Physics

PART II

DAVID HALLIDAY
ROBERT RESNICK

SELECTED PHYSICAL CONSTANTS

(See Appendix A for a more complete list)

Speed of light	c	3.00×10^8 meters/sec $= 1.86 \times 10^5$ miles/sec
Mass-energy relation	$c^2 (= E/m)$	931 Mev/amu $= 8.99 \times 10^{16}$ joules/kg
Gravitational constant	G	6.67×10^{-11} nt-m^2/kg^2
Universal gas constant	R	8.31 joules/mole °K = 1.99 cal/mole °K = 0.0823 li-atm/mole °K
Triple point of water	T_{tr}	273.16 °K
Permeability constant	μ_0	1.26×10^{-6} henry/meter
Permittivity constant	ϵ_0	8.85×10^{-12} farad/meter
Avogadro's constant	N_0	6.02×10^{23} molecules/mole
Boltzmann's constant	k	1.38×10^{-23} joule/molecule °K
Planck's constant	h	6.63×10^{-34} joule-sec
Elementary charge	e	1.60×10^{-19} coul
Electron rest mass	m_e	9.11×10^{-31} kg
Electron charge-to-mass ratio	e/m_e	1.76×10^{11} coul/kg
Proton rest mass	m_p	1.67×10^{-27} kg
Electron magnetic moment	μ_e	9.27×10^{-24} joule/tesla

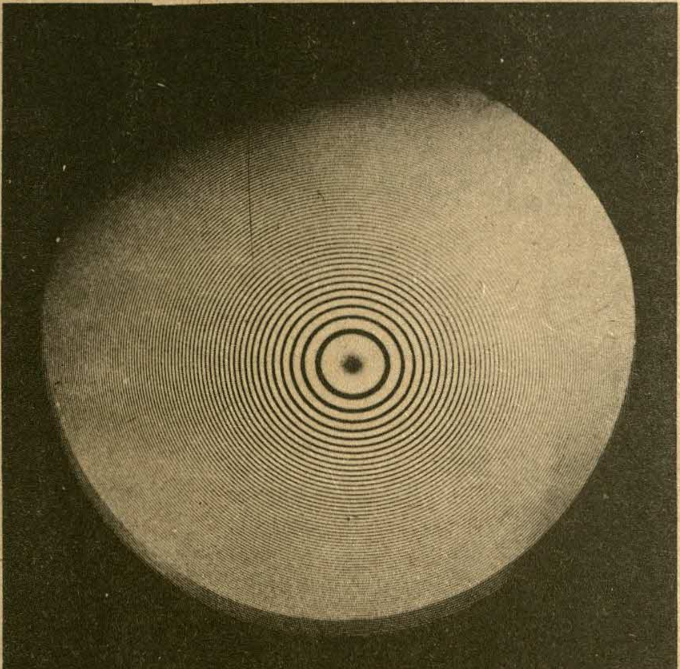
SELECTED PHYSICAL PROPERTIES

Density of air (STP)	1.29 kg/meter^3
Density of water (20° C)	1.00×10^3 kg/meter^3
Density of mercury (20° C)	13.5×10^3 kg/meter^3
Speed of sound in dry air (STP)	331 meters/sec = 1090 ft/sec
Acceleration of gravity (standard)	9.81 meters/sec^2 = 32.2 ft/sec^2
Standard atmosphere	1.01×10^5 nt/meter^2 = 14.7 lb/in.^{2} = 760 mm-Hg
Mean radius of the earth	6.37×10^6 meters = 3960 miles
Mean earth-sun distance	1.49×10^8 km = 92.9×10^6 miles
Mean earth-moon distance	3.80×10^5 km = 2.39×10^5 miles
Mass of earth	5.98×10^{24} kg
Heat of fusion of water (0° C, 1 atm)	79.7 cal/gm
Heat of vaporization of water (100° C, 1 atm)	539 cal/gm
Melting point of ice	0.00° C = 273.15° K
Ratio of specific heats (γ) for air (20° C)	1.40
Wavelength of the sodium yellow doublet	5892 A
Index of refraction of water (@ 5892 A)	1.33
Index of refraction of crown glass (@ 5892 A)	1.52

2060

Physics PART II

Wiley Eastern University Edition



245

UBSPD
UBS Publishers' Distributors Ltd
8/1-B, Chowringhee Lane, Calcutta-700 016
Ph:212451, 249473

Physics

PART II

DAVID HALLIDAY
Professor of Physics
University of Pittsburgh

ROBERT RESNICK
Professor of Physics
Rensselaer Polytechnic Institute

WILEY EASTERN LIMITED
New Delhi Bangalore Bombay Calcutta Madras Hyderabad

Books by Halliday (D.)
    Introductory Nuclear Physics, Second Edition
Books By Halliday (D.) and Resnick (R.)
    Physics for Students of Science and Engineering
    Parts I, First Edition and II, Second Edition Combined
    Physics, Part II
    Physics, Part I and II Combined
    Fundamentals of Physics
Books by Resnick (R.)
    Introduction to Special Relativity
Books by Resnick (R.) and Halliday (D.)
    Physics, Part I

First Edition, 1960
Second Edition, 1962
Revised Printing, 1966
Twenty Eighth Wiley Eastern Reprint, March 1989

Authorized reprint of the edition published by
John Wiley & Sons, Inc. New York, Chichester, Brisbane and Toronto
Copyright © 1962, 1966, John Wiley & Sons, Inc.
All rights reserved. No part of this book may be reproduced in any form
without the written permission of Wiley-Interscience, Inc.

This book is not to be sold outside the country
to which it is consigned by Wiley Eastern Limited

Sales Area : India, Sri Lanka and Bangladesh

30 29 28

ACC NO — 16335

Rs. 45.00

ISBN 0 85226 355 4

Published by V.R. Damodaran for Wiley Eastern Limited, 4835/24 Ansari Road, Daryaganj, New Delhi 110 002 and Printed by Gajendra Singh at Jugnu Offset, 1/11828 Panchsheel Garden, Naveen Shahdara, Delhi - 110 032.
Printed in India.

Physics, Part II, is a revised printing of the second edition of *Physics for Students of Science and Engineering*. In this (1966) new printing of Part II, the following revisions have been made.

1. A supplementary set of 142 problems has been provided. These have been carefully assembled so as to permit, when combined with the numerous problems in the original set, a wider choice of area of interest and level of difficulty than heretofore available.

2. A new Topical Supplement, on the differential form of Maxwell's equations and the electromagnetic wave equation, has been written for reference of students whose instructors choose to discuss this more advanced material.

3. The original appendices have been expanded and modernized and some new ones have been added in order to increase the usefulness and accuracy of such reference material. Front and back end-papers have been changed accordingly

4. Pagination, numbered references, symbols and units, have been changed, where necessary, for consistency with the new Part I and the above revisions of Part II.

Preface to the Second Edition

This revision of Part II of *Physics for Students of Science and Engineering*, which we hope will make our book more useful to students and instructors, is based on classroom experience at many institutions during the last two years. The basic outline of the book and its underlying philosophy remain unchanged. We have, however, modified the treatment of many topics, among them Ampère's, Faraday's, Gauss's and Lenz's laws and Huygens' principle, to make them clearer and more explicit than before. The experimental basis of the laws of physics has received increased stress, and new experimental material, such as the radiation pressure measurements of Nichols and Hull, optical masers, relativistic particle accelerators, and fiber optics illustrate the basic concepts. More than half the figures have been redrawn for greater clarity, and new figures and photographs have been inserted in key places. The questions and problems have been re-evaluated, and more than 200 are new.

In a few places we have supplied new sections either to extend or to improve the understanding of the basic material. For example, there are sections in reduced type on the electric vectors (**E**, **D**, and **P**) and on the magnetic vectors (**B**, **H**, and **M**). *RC* circuits are discussed, the Compton effect has been included in the treatment of the wave-particle duality, and a discussion of electromagnetic fields as seen by observers in relative motion has been added.

Other significant changes are (1) references to the literature available to students have been supplemented and brought up to date; (2) the sign conventions for geometrical optics have been put on a more physical basis; (3) attention has been called to examples or problems that contain important new material by using italicized titles; and (4) we have further systemized and simplified the notation in several places, particularly in wave optics.

We continue to be indebted to Dr. Benjamin Chi for assistance in preparing the illustrations and for his detailed criticism of the optics chapters. One of us (DH) wishes to thank Professor G. Wataghin for courtesies extended at the University of Turin while this revision was in progress. Finally, we extend special thanks to those students and instructors who have been kind enough to write to us about their experiences with the first edition.

DAVID HALLIDAY
ROBERT RESNICK

January 1, 1962
Pittsburgh, Pennsylvania
Troy, New York

Preface to First Edition

The time lag between developments in basic science and their application to engineering practice has shrunk enormously in the past few decades. The base of engineering, once largely empirical, is now largely scientific. Today the need is to stress principles rather than specific procedures, to select areas of contemporary interest rather than of past interest, and to condition the student to the atmosphere of change he will encounter during his career. These developments require a revision of the traditional course in general physics for engineers and scientists.

The most frequent criticisms made in varying degrees of textbooks used in such a course are these: (*a*) the content is encyclopedic in that topics are not treated with sufficient depth, the discussions are largely descriptive rather than explanatory and analytical, and too many topics are surveyed; (*b*) the content is not sufficiently "modern," and applications are drawn mostly from past engineering practice rather than from contemporary physics; (*c*) the organization of the material is too compartmentalized to reveal the essential unity of physics and its principles; (*d*) the approach is highly deductive and does not stress sufficiently the connection between theory and experiment. Of course, it is unlikely that a textbook will ever be written that is not criticized on one ground or another.

In writing this textbook we have been cognizant of these criticisms and have given much thought to ways of meeting them. We have considered the possibility of reorganizing the subject matter. The adoption of an atomic approach from the beginning or a structure built around energy in its various aspects suggest themselves. We have concluded that our goals can best be achieved by modifying the selection and treatment of topics within the traditional organization. To shuffle freely the cards of subject matter content or to abandon entirely a sequence which represents the growth of physical thought invites both a failure to appreciate the Newtonian and Maxwellian syntheses of classical physics and a superficial understanding of modern physics. A solid underpinning of classical physics is essential to build the superstructure of contemporary physics in our opinion.

To illustrate how we hope to achieve our goals within this framework, we present here the principal features of our book.

1. Many topics are treated in greater depth than has been customary heretofore, and much contemporary material has been woven into the body of the text. For example, gravitation, kinetic theory, electromagnetic waves, and physical optics, among others, are treated in greater depth. Contemporary topics, such as atomic standards, collision cross section, intermolecular forces, mass-energy conversion, isotope separation, the Hall effect, the free-electron model of conductivity, nuclear stability, nuclear resonance, and neutron diffraction, are discussed where they are pertinent.

To permit this greater depth and inclusion of contemporary material, we have omitted entirely or treated only indirectly much traditional material, such as simple machines, surface tension, viscosity, calorimetry, change of state, humidity, pumps, practical engines, musical scales, architectural acoustics, electrochemistry, thermoelectricity, motors, alternating-current circuits, electronics, lens aberrations, color, photometry, and others.

2. We have tried to reveal the unity of physics in many ways. Throughout the book we stress the general nature of key ideas common to all areas of physics. For example, the conservation laws of energy, linear momentum, angular momentum, and charge are used repeatedly. Wave concepts and properties of vibrating systems, such as resonance, are used in mechanics, sound, electromagnetism, optics, atomic physics, and nuclear physics. The field concept is applied to gravitation, fluid flow, electromagnetism, and nuclear physics.

The interrelation of the various disciplines of physics is emphasized by the use of physical and mathematical analogies and by similarity of method. For example, the correspondences between the mass-spring system and the *LC* circuit or between the acoustic tube and the electromagnetic cavity are emphasized, and the interweaving of microscopic and macroscopic approaches is noted in heat phenomena and electrical and magnetic phenomena. We have tried to make a smooth transition between particle mechanics and kinetic theory, stressing that, in their classical aspects, both belong to the Newtonian synthesis. We have also sought a smooth transition between electromagnetism and wave optics, pointing frequently to the Maxwellian synthesis.

We discuss the limitations of classical ideas and the domain of their validity, and we emphasize the generalizing nature of contemporary ideas applicable in a broader domain. Throughout we aim to show the relation of theory to experiment and to develop an awareness of the nature and uses of theory.

3. Our approach to quantum physics is not the traditional descriptive one. Rather we seek to develop the contemporary concepts fairly rigorously, at a length and depth appropriate to an introductory course. In the early chapters we pave the way by pointing to the limitations of classical theory, by stressing the aspects of classical physics that bear on contemporary physics, and by choosing illustrative examples that have a modern flavor. Thus we stress fields rather than circuits, particles rather than extended bodies, and wave optics rather than geometrical optics. Among the illustrative examples are molecular potential energy curves, binding energy of a deuteron, nuclear collisions, the nuclear model of the atom, the Thomson atom model, molecular dipoles, drift speed of electrons, stability of betatron orbits, nuclear magnetic resonance, the red shift, and others too numerous to mention.

The point of view is that of developing the fundamental ideas of quantum physics. The customary descriptive chapter on nuclear physics is, for example, not present. Instead, the wave-particle duality, the uncertainty principle, the complementarity principle, and the correspondence principle are stressed.

4. The mathematical level of our book assumes a concurrent course in calculus. The derivative is introduced in Chapter 3 and the integral in Chapter 7. The related physical concepts of slope and area under a curve are developed steadily. Calculus is used freely in the latter half of the book. Simple differential equations are not avoided, although no formal procedures are needed or given for solving them. Vector notation and vector algebra, including scalar and vector products, are used throughout. Displacement is taken as the prototype vector, and the idea of invariance of vector relations is developed.

5. The number of problems is unusually large, but few are "plug-in" problems. Many involve extensions of the text material, contemporary applications, or derivations. The questions at the end of each chapter are intended to be thought-provoking; they may serve as the basis for class discussion, for essay papers, or for self-study. Only rarely can the questions be answered by direct quotation from the text.

6. The book contains an unusually large number of worked-out examples, with the "plug-in" variety used only to emphasize a numerical magnitude. Algebraic, rather than numerical, solutions are stressed. Examples sometimes extend the text treatment or discuss the fine points, but usually they are applications of the principles, often of contemporary physics.

7. The textbook has been designed to fit physics courses of various lengths. In small print there is a great deal of supplementary material of an advanced, historical, or philosophical character, to be omitted or included to varying degrees depending on interest and course length. In addition, many chapters may be regarded as optional. Each teacher will make his own choice. At our institutions Chapter 14 (statics of rigid bodies) and Chapters 41 and 42 (geometrical optics) are omitted. Other possibilities suggested, depending on emphasis or depth desired, or the nature of succeeding courses, are Chapter 12 (rotational dynamics), Chapters 17 and 18 (fluids), Chapter 24 (kinetic theory—II), Chapter 32 (emf and circuits), Chapter 46 (polarization), and Chapters 47 and 48 (quantum physics).

8. We have adopted the mks system of units throughout, although the British engineering system is also used in mechanics. Having observed the gradual exclusion, year by year, of the cgs system from advanced textbooks, we have seen fit to limit ourselves to the bare definition of the basic cgs quantities. An extensive list of conversion factors appears in Appendix G.

We wish to thank the engineering and science students at both Rensselaer Polytechnic Institute and the University of Pittsburgh who have borne with us through two successive preliminary editions. Constructive criticisms from our colleagues at each institution and from some eight reviewers have resulted in many changes. Benjamin Chi of R.P.I. has been of major service in all aspects of the preparation of the manuscript. Finally, we express our deep appreciation to our wives, not only for aid in typing and proofreading but for the patience and encouragement without which this book might never have been written.

DAVID HALLIDAY
ROBERT RESNICK

January 1960
Pittsburgh, Pennsylvania
Troy, New York

Contents

26 **CHARGE AND MATTER 647**

26–1 Electromagnetism—A Preview 647
26–2 Electric Charge 648
26–3 Conductors and Insulators 649
26–4 Coulomb's Law 650
26–5 Charge Is Quantized 654
26–6 Charge and Matter 655
26–7 Charge Is Conserved 660

27 **THE ELECTRIC FIELD 663**

27–1 The Electric Field 663
27–2 The Electric Field Strength **E** 665
27–3 Lines of Force 665
27–4 Calculation of **E** 670
27–5 A Point Charge in an Electric Field 675
27–6 A Dipole in an Electric Field 678

28 **GAUSS'S LAW 686**

28–1 Flux of the Electric Field 686
28–2 Gauss's Law 690
28–3 Gauss's Law and Coulomb's Law 690
28–4 An Insulated Conductor 692
28–5 Experimental Proof of Gauss's and Coulomb's Laws 693
28–6 Gauss's Law—Some Applications 695
28–7 The Nuclear Model of the Atom 700

29 ELECTRIC POTENTIAL 708

29–1 Electric Potential 708
29–2 Potential and Field Strength 711
29–3 Potential Due to a Point Charge 715
29–4 A Group of Point Charges 716
29–5 Potential Due to a Dipole 718
29–6 Electric Potential Energy 722
29–7 Calculation of E from V 724
29–8 An Insulated Conductor 730
29–9 The Electrostatic Generator 733

30 CAPACITORS AND DIELECTRICS 741

30–1 Capacitance 741
30–2 Calculating Capacitance 745
30–3 Paralle-Plate Capacitor with Dielectric 748
30–4 Dielectrics—An Atomic View 750
30–5 Dielectrics and Gauss's Law 753
30–6 Three Electric Vectors 756
30–7 Energy Storage in an Electric Field 760

31 CURRENT AND RESISTANCE 770

31–1 Current and Current Density 770
31–2 Resistance, Resistivity, and Conductivity 774
31–3 Ohm's Law 778
31–4 Resistivity—An Atomic View 780
31–5 Energy Transfers in an Electric Circuit 783

32 ELECTROMOTIVE FORCE AND CIRCUITS 789

32–1 Electromotive Force 789
32–2 Calculating the Current 792
32–3 Other Single-Loop Circuits 793
32–4 Potential Differences 795
32–5 Multiloop Circuits 798
32–6 Measuring Currents and Potential Differences 800
32–7 The Potentiometer 801
32–8 *RC* Circuits 802

33 THE MAGNETIC FIELD 814

33–1 The Magnetic Field 814
33–2 The Definition of **B** 816
33–3 Magnetic Force on a Current 819
33–4 Torque on a Current Loop 823
33–5 The Hall Effect 827
33–6 Circulating Charges 829
33–7 The Cyclotron 831
33–8 Thomson's Experiment 835

34 AMPERE'S LAW 844

34–1 Ampère's Law 844
34–2 **B** Near a Long Wire 848
34–3 Magnetic Lines of Induction 851
34–4 Two Parallel Conductors 853
34–5 **B** for a Solenoid 856
34–6 The Biot-Savart Law 859

35 FARADAY'S LAW 870

35–1 Faraday's Experiments 870
35–2 Faraday's Law of Induction 872
35–3 Lenz's Law 873
35–4 Induction—A Quantitative Study 875
35–5 Time-Varying Magnetic Fields 880
35–6 The Betatron 885
35–7 Induction and Relative Motion 888

36 INDUCTANCE 899

36–1 Inductance 899
36–2 Calculation of Inductance 901
36–3 An *LR* Circuit 902
36–4 Energy and the Magnetic Field 907
36–5 Energy Density and the Magnetic Field 910

37 MAGNETIC PROPERTIES OF MATTER 915

37–1 Poles and Dipoles 915
37–2 Gauss's Law for Magnetism 920
37–3 Paramagnetism 920
37–4 Diamagnetism 923
37–5 Ferromagnetism 926
37–6 Nuclear Magnetism 931
37–7 Three Magnetic Vectors 934

38 ELECTROMAGNETIC OSCILLATIONS 943

38–1 *LC* Oscillations 943
38–2 Analogy to Simple Harmonic Motion 947
38–3 Electromagnetic Oscillations—Quantitative 948
38–4 Forced Oscillations and Resonance 952
38–5 Lumped and Distributed Elements 954
38–6 Electromagnetic Cavity Oscillator 955
38–7 Induced Magnetic Fields 959
38–8 Displacement Current 962
38–9 Maxwell's Equations 963
38–10 Maxwell's Equations and Cavity Oscillations 963

39 ELECTROMAGNETIC WAVES 971

39–1 Transmission Lines 971
39–2 Coaxial Cable—Fields and Currents 974
39–3 Waveguide 977
39–4 Radiation 979
39–5 Traveling Waves and Maxwell's Equations 982
39–6 The Poynting Vector 986

40 NATURE AND PROPAGATION OF LIGHT 993

40–1 Light and the Electromagnetic Spectrum 993
40–2 Energy and Momentum 995
40–3 The Speed of Light 998
40–4 Moving Sources and Observers 1003
40–5 Doppler Effect 1006

41 REFLECTION AND REFRACTION—PLANE WAVES AND PLANE SURFACES 1013

41–1 Reflection and Refraction 1013
41–2 Huygens' Principle 1018
41–3 Huygens' Principle and the Law of Reflection 1020
41–4 Huygens' Principle and the Law of Refraction 1021
41–5 Total Internal Reflection 1024
41–6 Fermat's Principle 1028

42 REFLECTION AND REFRACTION—SPHERICAL WAVES AND SPHERICAL SURFACES 1035

42–1 Geometrical Optics and Wave Optics 1035
42–2 Spherical Waves—Plane Mirror 1038
42–3 Spherical Waves—Spherical Mirror 1041
42–4 Spherical Refracting Surface 1049
42–5 Thin Lenses 1055

43 INTERFERENCE 1068

43–1 Young's Experiment 1068
43–2 Coherence 1074
43–3 Intensity in Young's Experiment 1077
43–4 Adding Wave Disturbances 1080
43–5 Interference from Thin Films 1083
43–6 Phase Changes on Reflection 1089
43–7 Michelson's Interferometer 1090
43–8 Michelson's Interferometer and Light Propagation 1092

44 DIFFRACTION 1099

44–1 Introduction 1099
44–2 Single Slit 1102
44–3 Single Slit—Qualitative 1105

44-4 Single Slit—Quantitative 1108
44-5 Diffraction at a Circular Aperture 1112
44-6 Double Slit 1115

45 GRATINGS AND SPECTRA 1123

45-1 Introduction 1123
45-2 Multiple Slits 1123
45-3 Diffraction Gratings 1128
45-4 Resolving Power of a Grating 1132
45-5 X-ray Diffraction 1134
45-6 Bragg's Law 1140

46 POLARIZATION 1147

46-1 Polarization 1147
46-2 Polarizing Sheets 1150
46-3 Polarization by Reflection 1153
46-4 Double Refraction 1155
46-5 Circular Polarization 1163
46-6 Angular Momentum of Light 1167
46-7 Scattering of Light 1167
46-8 Double Scattering 1169

47 LIGHT AND QUANTUM PHYSICS 1173

47-1 Sources of Light 1173
47-2 Cavity Radiators 1174
47-3 Planck's Radiation Formula 1177
47-4 Photoelectric Effect 1179
47-5 Einstein's Photon Theory 1182
47-6 The Compton Effect 1184
47-7 Line Spectra 1188
47-8 The Hydrogen Atom 1191
47-9 The Correspondence Principle 1196

48 WAVES AND PARTICLES 1200

48-1 Matter Waves 1200
48-2 Atomic Structure and Standing Waves 1203
48-3 Wave Mechanics 1204
48-4 Meaning of Ψ 1208
48-5 The Uncertainty Principle 1210

SUPPLEMENTARY TOPIC V 1

The Differential Form of Maxwell's Equations and the Electromagnetic Wave Equation

SUPPLEMENTARY PROBLEMS 9

APPENDICES 29

- **A** Definition of Standards and Fundamental and Derived Physical Constants 29
- **B** Miscellaneous Terrestrial Data 32
- **C** The Solar System 34
- **D** Periodic Table of the Elements 36
- **E** The Particles of Physics 37
- **F** Symbols, Dimensions, and Units for Physical Quantities 39
- **G** Conversion Factors 41
- **H** Mathematical Symbols and the Greek Alphabet 49
- **I** Mathematical Formulas 52
- **J** Values of Trigonometric Functions 53
- **K** Nobel Prize Winners in Physics 55
- **L** The Gaussian System of Units 58

ANSWERS TO PROBLEMS 61

ANSWERS TO SUPPLEMENTARY PROBLEMS 68

INDEX 73

Physics PART II

Charge and Matter

CHAPTER 26

26–1 Electromagnetism—A Preview

The science of electricity has its roots in the observation, known to Thales of Miletus in 600 B.C., that a rubbed piece of amber will attract bits of straw. The study of magnetism goes back to the observation that naturally occurring "stones" (that is, magnetite) will attract iron. These two sciences developed quite separately until 1820, when Hans Christian Oersted (1777–1851) observed a connection between them, namely, that an *electric* current in a wire can affect a *magnetic* compass needle (Section 33–1).

The new science of electromagnetism was developed further by many workers, of whom one of the most important was Michael Faraday (1791–1867). It fell to James Clerk Maxwell (1831–1879) to put the laws of electromagnetism in the form in which we know them today. These laws, often called *Maxwell's equations*, are displayed in Table 38–3, which the student may want to examine at this time. These laws play the same role in electromagnetism that Newton's laws of motion and of gravitation do in mechanics.

Although Maxwell's synthesis of electromagnetism rests heavily on the work of his predecessors, his own contribution is central and vital. Maxwell deduced that light is electromagnetic in nature and that its speed can be found by making purely electric and magnetic measurements. Thus the science of optics was intimately connected with those of electricity and of magnetism. The scope of Maxwell's equations is remarkable, including as it does the fundamental principles of all large-scale electromagnetic and optical devices such as motors, cyclotrons, electronic computers, radio, television, microwave radar, microscopes, and telescopes.

The development of classical electromagnetism did not end with Maxwell. The English physicist Oliver Heaviside (1850–1925) and especially the Dutch physicist H. A. Lorentz (1853–1928) contributed substantially to the clarification of Maxwell's theory. Heinrich Hertz (1857–1894) * took a great step forward when, more than twenty years after Maxwell set up his theory, he produced in the laboratory electromagnetic "Maxwellian waves" of a kind that we would now call short radio waves. It remained for Marconi and others to exploit this practical application of the electromagnetic waves of Maxwell and Hertz.

Present interest in electromagnetism takes two forms. At the level of engineering applications Maxwell's equations are used constantly and universally in the solution of a wide variety of practical problems. At the level of the foundations of the theory there is a continuing effort to extend its scope in such a way that electromagnetism is revealed as a special case of a more general theory. Such a theory would also include (say) the theories of gravitation and of quantum physics. This grand synthesis has not yet been achieved.

26–2 Electric Charge

The rest of this chapter deals with electric charge and its relationship to matter. We can show that there are *two kinds* of charge by rubbing a glass rod with silk and hanging it from a long silk thread as in Fig. 26–1. If a second glass rod is rubbed with silk and held near the rubbed end of the first rod, the rods will repel each other. On the other hand, a hard-rubber rod rubbed with fur will *attract* the glass rod. Two hard-rubber rods rubbed with fur will repel each other. We explain these facts by saying that rubbing a rod gives it an *electric charge* and that the charges on the two rods exert forces on each other. Clearly the charges on the glass and on the hard rubber must be different in nature.

Benjamin Franklin (1706–1790), who, among his other achievements, was the first American physicist, named the kind of electricity that appears on the glass *positive* and the kind that appears on the hard rubber *negative;* these names have remained to this day. We can sum up these experiments by saying that *like charges repel and unlike charges attract.*

Electric effects are not limited to glass rubbed with silk or to hard rubber rubbed with fur. Any substance rubbed with any other under suitable conditions will become charged to some extent; by comparing the unknown charge with a charged glass rod or a charged hard-rubber rod, it can be labeled as either positive or negative.

The modern view of bulk matter is that, in its normal or neutral state, it contains equal amounts of positive and negative electricity. If two bodies like glass and silk are rubbed together, a small amount of charge is transferred from one to the other, upsetting the electric neutrality of each. In this case the glass would become positive, the silk negative.

---

* "Heinrich Hertz," by P. and E. Morrison, *Scientific American*, December 1957.

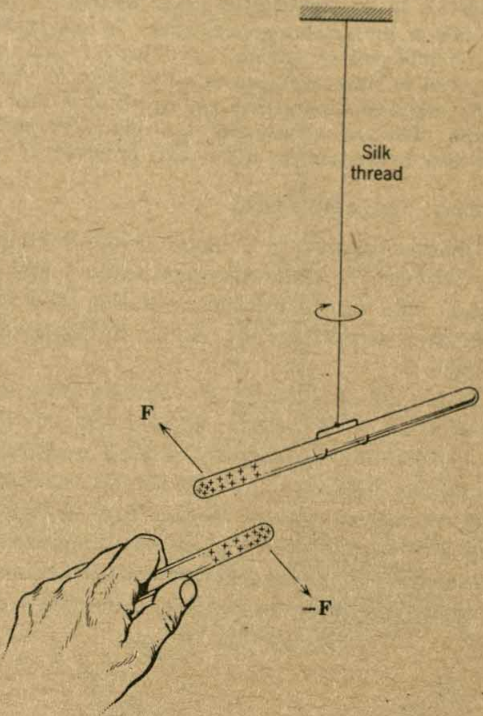


Fig. 26-1 Two positively charged glass rods repel each other.

26-3 Conductors and Insulators

A metal rod held in the hand and rubbed with fur will not seem to develop a charge. It is possible to charge such a rod, however, if it is furnished with a glass or hard-rubber handle and if the metal is not touched with the hands while rubbing it. The explanation is that metals, the human body, and the earth are *conductors* of electricity and that glass, hard rubber, plastics, etc., are *insulators* (also called *dielectrics*).

In conductors electric charges are free to move through the material, whereas in insulators they are not. Although there are no perfect insulators, the insulating ability of fused quartz is about 10^{25} times as great as that of copper, so that for many practical purposes some materials behave as if they were perfect insulators.

In metals a fairly subtle experiment called the *Hall effect* (see Section 33-5) shows that only negative charge is free to move. Positive charge is as immobile as it is in glass or in any other dielectric. The actual charge carriers in metals are the *free electrons*. When isolated atoms are combined to form a metallic solid, the outer electrons of the atom do not remain attached to individual atoms but become free to move throughout the volume of the solid. For some conductors, such as electrolytes, both positive and negative charges can move.

A class of materials called *semiconductors* is intermediate between conductors and insulators in its ability to conduct electricity. Among the elements, silicon and germanium are well-known examples. In semiconductors the electrical conductivity can often be greatly increased by adding very small amounts of other elements; traces of arsenic or boron are often added to silicon for this purpose. Semiconductors have many practical applications, among which is their use in the construction of transistors. The mode of action of semiconductors cannot be described adequately without some understanding of the basic principles of quantum physics.

26–4 Coulomb's Law

Charles Augustin de Coulomb (1736–1806) in 1785 first measured electrical attractions and repulsions quantitatively and deduced the law that governs them. His apparatus, shown in Fig. 26–2, resembles the hanging rod of Fig. 26–1, except that the charges in Fig. 26–2 are confined to small spheres a and b.

If a and b are charged, the electric force on a will tend to twist the suspension fiber. Coulomb canceled out this twisting effect by turning the suspension head through the angle θ needed to keep the two charges at the particular distance apart in which he was interested. The angle θ is then a relative measure of the electric force acting on charge a. The device of Fig. 26–2 is called a *torsion balance;* a similar arrangement was used later by Cavendish to measure gravitational attractions (Section 16–3).

Coulomb's first experimental results can be represented by

$$F \propto \frac{1}{r^2}.$$

F is the magnitude of the force that acts on each of the two charges a and b; r is their distance apart. These forces, as Newton's third law requires, act along the line joining the charges but point in opposite directions. Note that the magnitude of the force on each charge is the same, even though the charges may be different.

Coulomb also studied how the electrical force varied with the relative size of the charges on the spheres of his torsion balance. For example, if we touch a charged conducting sphere to an exactly similar but uncharged conducting sphere, the original charge must divide equally between the spheres. By such techniques Coulomb extended the inverse square relationship to

$$F \propto \frac{q_1 q_2}{r^2}, \tag{26–1}$$

where q_1 and q_2 are relative measures of the charges on spheres a and b. Equation 26–1, which is called *Coulomb's law,* holds only for charged objects whose sizes are much smaller than the distance between them. We often say that it holds only for *point charges.*

Coulomb's law resembles the inverse square law of gravitation which was already more than 100 years old at the time of Coulomb's experiments: q plays the role of m in that law. In gravity, however, the forces are always

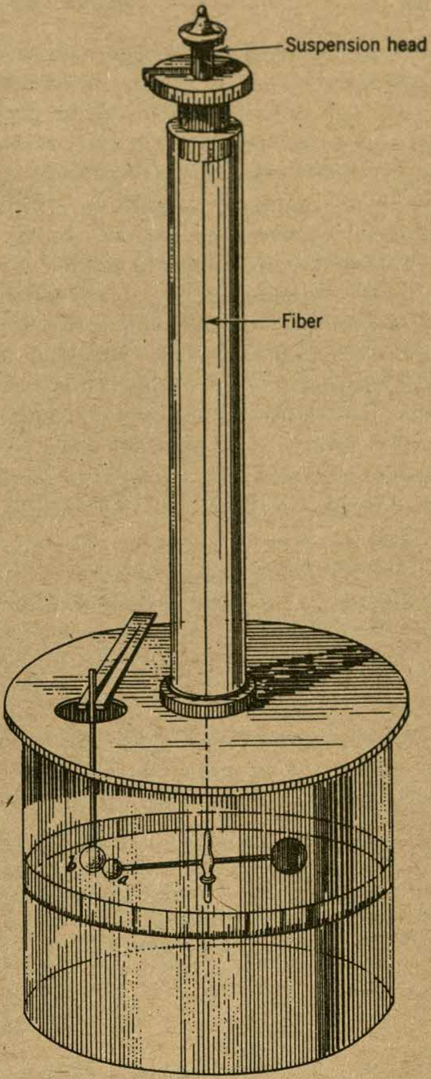


Fig. 26-2 Coulomb's torsion balance from his 1785 memoir to the French Academy of Sciences.

attractive; this corresponds to the fact that there are two kinds of electricity but (apparently) only one kind of mass.

Our belief in Coulomb's law does not rest quantitatively on Coulomb's experiments. Torsion balance measurements are difficult to make to an accuracy of better than a few per cent. Such measurements could not, for example, convince us that the exponent in Eq. 26-1 is exactly 2 and not, say, 2.01. In Section 28-5 we show that Coulomb's law can also be deduced from an indirect experiment which shows that the exponent in Eq. 26-1 lies between the limits of approximately 2.000000002 and 1.999999998. Small wonder that we usually assume the exponent to be exactly 2.

Although we have established the physical concept of electric charge, we have not yet defined a unit in which it may be measured. It is possible to do so operationally by putting equal charges q on the spheres of a torsion balance and by measuring the magnitude F of the force that acts on each when the charges are a measured distance r apart. One could then define q to have a unit value if a unit force acts on each charge when the charges are separated by a unit distance and one can give a name to the unit of charge so defined.*

For practical reasons having to do with the accuracy of measurements, the unit of charge in the mks system is not defined using a torsion balance but is derived from the unit of electric current. If the ends of a long wire are connected to the terminals of a battery, it is common knowledge that an *electric current* i is set up in the wire. We visualize this current as a flow of charge. The mks unit of current is the *ampere* (abbr. *amp*). In Section 34-4 we describe the operational procedures in terms of which the ampere is defined.

The mks unit of charge is the *coulomb* (abbr. *coul*). *A coulomb is defined as the amount of charge that flows through a given cross section of a wire in 1 second if there is a steady current of 1 ampere in the wire.* In symbols

$$q = it, \tag{26-2}$$

where q is in coulombs if i is in amperes and t is in seconds. Thus, if a wire is connected to an insulated metal sphere, a charge of 10^{-6} coul can be put on the sphere if a current of 1.0 amp exists in the wire for 10^{-6} sec.

▶ **Example 1.** A copper penny has a mass of 3.1 gm. Being electrically neutral, it contains equal amounts of positive and negative electricity. What is the magnitude q of these charges? A copper atom has a positive nuclear charge of 4.6×10^{-18} coul and a negative electronic charge of equal magnitude.

The number N of copper atoms in a penny is found from the ratio

$$\frac{N}{N_0} = \frac{m}{M},$$

where N_0 is Avogadro's number, m the mass of the coin, and M the atomic weight of copper. This yields, solving for N,

$$N = \frac{(6.0 \times 10^{23}\ \text{atoms/mole})(3.1\ \text{gm})}{64\ \text{gm/mole}} = 2.9 \times 10^{22}\ \text{atoms}.$$

The charge q is

$$q = (4.6 \times 10^{-18}\ \text{coul/atom})(2.9 \times 10^{22}\ \text{atoms}) = 1.3 \times 10^{5}\ \text{coul}.$$

In a 100-watt, 110-volt light bulb the current is 0.91 amp. The student should verify that it would take 40 hr for a charge of this amount to pass through this bulb. ◀

---

* This scheme is the basis for the definition of the unit of charge called the *statcoulomb*. However, in this book we do not use this unit or the systems of units of which it is a part; see Appendix L, however.

Equation 26–1 can be written as an equality by inserting a constant of proportionality. Instead of writing this simply as, say, k, it is usually written in a more complex way as $1/4\pi\epsilon_0$ or

$$F = \frac{1}{4\pi\epsilon_0}\frac{q_1q_2}{r^2}. \tag{26–3}$$

Certain equations that are derived from Eq. 26–3, but are used more often than it is, will be simpler in form if we do this.

In the mks system we can measure q_1, q_2, r, and F in Eq. 26–3 in ways that do not depend on Coulomb's law. Numbers with units can be assigned to them. There is no choice about the so-called *permittivity constant* ϵ_0; it must have that value which makes the right-hand side of Eq. 26–3 equal to the left-hand side. This (measured) value turns out to be *

$$\epsilon_0 = 8.85418 \times 10^{-12}\ \text{coul}^2/\text{nt-m}^2.$$

In this book the value 8.9×10^{-12} coul²/nt-m² will be accurate enough for all problems. For direct application of Coulomb's law or in any problem in which the quantity $1/4\pi\epsilon_0$ occurs we may use, with sufficient accuracy for this book,

$$1/4\pi\epsilon_0 = 9.0 \times 10^9\ \text{nt-m}^2/\text{coul}^2.$$

▶ **Example 2.** Let the total positive and the total negative charges in a copper penny be separated to a distance such that their force of attraction is 1.0 lb (= 4.5 nt). How far apart must they be?

We have (Eq. 26–3)

$$F = \frac{1}{4\pi\epsilon_0}\frac{q_1q_2}{r^2}.$$

Putting $q_1q_2 = q^2$ (see Example 1) and solving for r yields

$$r = q\sqrt{\frac{1/4\pi\epsilon_0}{F}} = 1.3 \times 10^5\ \text{coul}\sqrt{\frac{9.0 \times 10^9\ \text{nt-m}^2/\text{coul}^2}{4.5\ \text{nt}}}$$

$$= 5.8 \times 10^9\ \text{meters} = 3.6 \times 10^6\ \text{miles}.$$

This suggests that it is not possible to upset the electrical neutrality of gross objects by any very large amount. What would be the force between the two charges if they were placed 1.0 meter apart? ◀

If more than two charges are present, Eq. 26–3 holds for every pair of charges. Let the charges be q_1, q_2, and q_3, etc.; we calculate the force exerted on any one (say q_1) by all the others from the vector equation

$$\mathbf{F}_1 = \mathbf{F}_{12} + \mathbf{F}_{13} + \mathbf{F}_{14} + \cdots, \tag{26–4}$$

where $\mathbf{F}_{12}$, for example, is the force exerted on q_1 by q_2.

---

* For practical reasons this value is not actually measured by direct application of Eq. 26-3 but in an equivalent although more circuitous way that is described in Section 30-2.

▶ **Example 3.** Figure 26–3 shows three charges q_1, q_2, and q_3. What force acts on q_1? Assume that $q_1 = -1.0 \times 10^{-6}$ coul, $q_2 = +3.0 \times 10^{-6}$ coul, $q_3 = -2.0 \times 10^{-6}$ coul, $r_{12} = 15$ cm, $r_{13} = 10$ cm, and $\theta = 30°$.

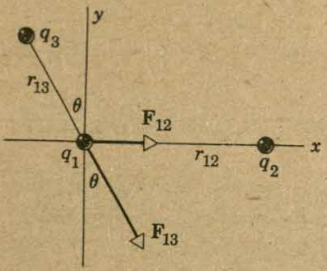


Fig. 26–3 Example 3. Showing the forces exerted on q_1 by q_2 and q_3.

From Eq. 26–3, ignoring the signs of the charges, since we are interested only in the magnitudes of the forces,

$$
\begin{aligned}
F_{12} &= \frac{1}{4\pi\epsilon_0}\frac{q_1q_2}{r_{12}^2} \\
&= \frac{(9.0 \times 10^9 \text{ nt-m}^2/\text{coul}^2)(1.0 \times 10^{-6}\text{ coul})(3.0 \times 10^{-6}\text{ coul})}{(1.5 \times 10^{-1}\text{ meter})^2} \\
&= 1.2 \text{ nt}
\end{aligned}
$$

and

$$
\begin{aligned}
F_{13} &= \frac{(9.0 \times 10^9 \text{ nt-m}^2/\text{coul}^2)(1.0 \times 10^{-6}\text{ coul})(2.0 \times 10^{-6}\text{ coul})}{(1.0 \times 10^{-1}\text{ meter})^2} \\
&= 1.8 \text{ nt}.
\end{aligned}
$$

The directions of $\mathbf{F}_{12}$ and $\mathbf{F}_{13}$ are as shown in the figure.

The components of the resultant force $\mathbf{F}_1$ acting on q_1 (see Eq. 26–4) are

$$
\begin{aligned}
F_{1x} &= F_{12x} + F_{13x} = F_{12} + F_{13}\sin\theta \\
&= 1.2\text{ nt} + (1.8\text{ nt})(\sin 30°) \doteq 2.1\text{ nt}
\end{aligned}
$$

and

$$
\begin{aligned}
F_{1y} &= F_{12y} + F_{13y} = 0 - F_{13}\cos\theta \\
&= -(1.8\text{ nt})(\cos 30°) = -1.6\text{ nt}.
\end{aligned}
$$

The student should find the magnitude of $\mathbf{F}_1$ and the angle it makes with the x-axis. ◀

26–5 Charge Is Quantized

In Franklin's day electric charge was thought of as a continuous fluid, an idea that was useful for many purposes. The atomic theory of matter, however, has shown that fluids themselves, such as water and air, are not continuous but are made up of atoms. Experiment shows that the "electric fluid" is not continuous either but that it is made up of integral multiples of a certain minimum electric charge. This fundamental charge, to which we give the symbol e, has the magnitude 1.60210×10^{-19} coul. Any physically existing charge q, no matter what its origin, can be written as ne where n is a positive or a negative integer.

When a physical property such as charge exists in discrete "packets" rather than in continuous amounts, the property is said to be *quantized*. Quantization is basic to modern physics. The existence of atoms and of particles like the electron and the proton indicates that *mass* is quantized also. Later the student will learn that several other properties prove to be quantized when suitably examined on the atomic scale; among them are energy and angular momentum.

The *quantum of charge* e is so small that the "graininess" of electricity does not show up in large-scale experiments, just as we do not realize that the air we breathe is made up of atoms. In an ordinary 110-volt, 100-watt light bulb, for example, 6×10^{18} elementary charges enter and leave the filament every second.

There exists today no theory that predicts the quantization of charge (or the quantization of mass, that is, the existence of fundamental particles such as protons, electrons, pions, etc.). Even assuming quantization, the classical theory of electromagnetism and Newtonian mechanics are incomplete in that they do not correctly describe the behavior of charge and matter on the atomic scale. The classical theory of electromagnetism, for example, describes correctly what happens when a bar magnet is thrust through a closed copper loop; it fails, however, if we wish to explain the magnetism of the bar in terms of the atoms that make it up. The more detailed theories of quantum physics are needed for this and similar problems.

26-6 Charge and Matter

Matter as we ordinarily experience it can be regarded as composed of three kinds of elementary particles, the proton, the neutron, and the electron. Table 26-1 shows their masses and charges. Note that the masses of the neutron and the proton are approximately equal but that the electron is lighter by a factor of about 1840.

Atoms are made up of a dense, positively charged *nucleus*, surrounded by a

Table 26-1

PROPERTIES OF THE PROTON, THE NEUTRON, AND THE ELECTRON

Particle	Symbol	Charge	Mass
Proton	p	$+e$	1.67252×10^{-27} kg
Neutron	n	0	1.67482×10^{-27} kg
Electron	e^-	$-e$	9.1091×10^{-31} kg

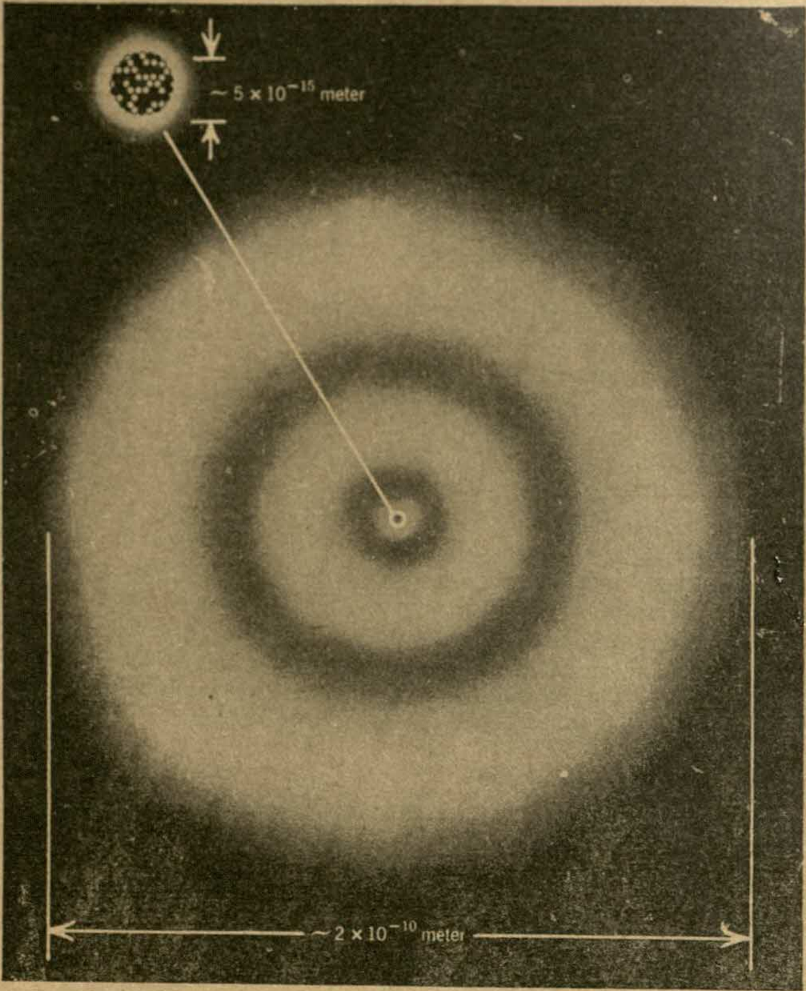


Fig. 26–4 An atom, suggesting the electron cloud and, above, an enlarged view of the nucleus.

cloud of electrons; see Fig. 26–4. The nucleus varies in radius from about 1×10^{-15} meter for hydrogen to about 7×10^{-15} meter for the heaviest atoms. The outer diameter of the electron cloud, that is, the diameter of the the atom, lies in the range 1–3 $\times 10^{-10}$ meter, about 10^5 times larger than the nuclear diameter.

▶ **Example 4.** The distance r between the electron and the proton in the hydrogen atom is about 5.3×10^{-11} meter. What are the magnitudes of (a) the electrical force and (b) the gravitational force between these two particles?

From Coulomb's law,

$$F_e = \frac{1}{4\pi\epsilon_0}\frac{q_1q_2}{r^2}$$

$$= \frac{(9.0 \times 10^{9}\ \text{nt-m}^2/\text{coul}^2)(1.6 \times 10^{-19}\ \text{coul})^2}{(5.3 \times 10^{-11}\ \text{meter})^2}$$

$$= 8.1 \times 10^{-8}\ \text{nt.}$$

The gravitational force is given by Eq. 16-1, or

$$F_g = G\frac{m_1m_2}{r^2}$$

$$= \frac{(6.7 \times 10^{-11}\ \text{nt-m}^2/\text{kg}^2)(9.1 \times 10^{-31}\ \text{kg})(1.7 \times 10^{-27}\ \text{kg})}{(5.3 \times 10^{-11}\ \text{meter})^2}$$

$$= 3.7 \times 10^{-47}\ \text{nt.}$$

Thus the electrical force is about 10^{39} times stronger than the gravitational force. ◀

The significance of Coulomb's law goes far beyond the description of the forces acting between charged balls or rods. This law, when incorporated into the structure of quantum physics, correctly describes (*a*) the electric forces that bind the electrons of an atom to its nucleus, (*b*) the forces that bind atoms together to form molecules, and (*c*) the forces that bind atoms or molecules together to form solids or liquids. Thus most of the forces of our daily experience that are not gravitational in nature are electrical. A force transmitted by a steel cable is basically an electrical force because, if we pass an imaginary plane through the cable at right angles to it, it is only the attractive electrical interatomic forces acting between atoms on opposite sides of the plane that keep the cable from parting. We ourselves are an assembly of nuclei and electrons bound together in a stable configuration by Coulomb forces.

In the atomic *nucleus* we encounter a new force which is neither gravitational nor electrical in nature. This strong attractive force, which binds together the protons and neutrons that make up the nucleus, is called simply *the nuclear force*. If this force were not present, the nucleus would fly apart at once because of the strong Coulomb repulsion force that acts between its protons. The nature of the nuclear force is only partially understood today and forms the central problem of present-day researches in nuclear physics.

▶ **Example 5.** What repulsive Coulomb force exists between two protons in a nucleus of iron? Assume a separation of 4.0×10^{-15} meter.

From Coulomb's law,

$$F = \frac{1}{4\pi\epsilon_0}\frac{q_1q_2}{r^2}$$

$$= \frac{(9.0 \times 10^{9}\ \text{nt-m}^2/\text{coul}^2)(1.6 \times 10^{-19}\ \text{coul})^2}{(4.0 \times 10^{-15}\ \text{meter})^2}$$

$$= 14\ \text{nt.}$$

This enormous repulsive force must be more than compensated for by the strong attractive nuclear forces. This example, combined with Example 4, shows that nuclear binding forces are much stronger than atomic binding forces. Atomic binding forces are, in turn, much stronger than gravitational forces for the same particles separated by the same distance. ◀

The repulsive Coulomb forces acting between the protons in a nucleus make the nucleus less stable than it otherwise would be. The spontaneous emission of alpha particles from heavy nuclei and the phenomenon of nuclear fission are evidences of this instability.

The fact that heavy nuclei contain significantly more neutrons than protons is still another effect of the Coulomb forces. Consider Fig. 26–5 in which a particular atomic species is represented by a circle, the coordinates being Z, the number of protons in the nucleus (that is, the *atomic number*), and N, the number of neutrons in the nucleus (that is, the *neutron number*). Stable nuclei are represented by filled circles and radioactive nuclei, that is, nuclei that disintegrate spontaneously, emitting electrons or α-particles, by open circles. Note that all elements (iron, for example, for which $Z = 26$; see arrow) exist in a number of different forms, called *isotopes*.

Figure 26–5 shows that light nuclei, for which the Coulomb forces are relatively unimportant,* lie on or close to the line labeled "$N = Z$" and thus have about equal numbers of neutrons and protons. The heavier nuclei have a pronounced neutron excess, U^{238} having 92 protons and 238 − 92 or 146 neutrons.† In the absence of Coulomb forces we would assume, extending the $N = Z$ rule, that the most stable nucleus with 238 particles would have 119 protons and 119 neutrons. However, such a nucleus, if assembled, would fly apart at once because of Coulomb repulsion. Relative stability is found only if 27 of the protons are replaced by neutrons, thus diluting the total Coulomb repulsion. Even in U^{238} Coulomb repulsion is still very important because (*a*) this nucleus is radioactive and emits α-particles, and (*b*) it may break up into two large fragments (*fission*) if it absorbs a neutron; both processes result in separation of the nuclear charge and are Coulomb repulsion effects. Figure 26–5 shows that *all* nuclei with $Z > 83$ are unstable.

We have pointed out that matter, as we ordinarily experience it, is made up of electrons, neutrons, and protons. Nature exhibits much more variety than this, however. No fewer than 28 distinct elementary particles are now known, most of them having been discovered since 1940, either in the penetrating cosmic rays that come to us from beyond our atmosphere or in the reaction products of giant cyclotron-like devices.

Appendix E, which lists some properties of these particles, shows that, like the more familiar particles of Table 26–1, their charges are quantized, the quantum of charge again being e. An understanding of the nature of these particles and of their relationships to each other is perhaps the most significant research goal of modern physics.

---

* Coulomb forces are important in relation to the strong nuclear attractive forces only for large nuclei, because Coulomb repulsion occurs between *every pair* of protons in the nucleus but the attractive nuclear force does not. In U^{238}, for example, every proton exerts a force of repulsion on each of the other 91 protons. However, each proton (and neutron) exerts a nuclear attraction on only a small number of other neutrons and protons that happen to be near it. As we proceed to larger nuclei, the amount of energy associated with the repulsive Coulomb forces increases much faster than that associated with the attractive nuclear forces.

† The superscript in this notation is the *mass number* A $(= N + Z)$. This is the total number of particles in the nucleus.

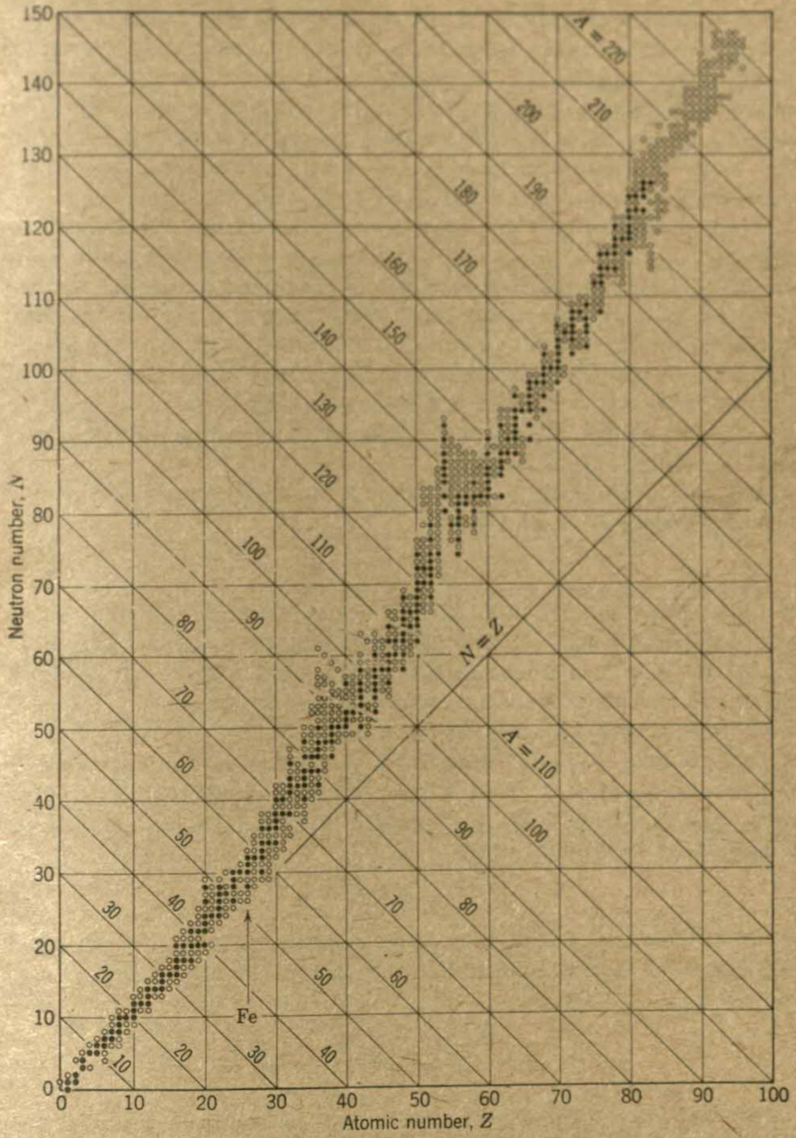


Fig. 26–5 The filled circles represent stable nuclei, the open ones radioactive nuclei. This figure does not show the more recently discovered radioactive nuclei, which appear hroughout the chart, extending up to $Z = 103$.

26-7 Charge Is Conserved

When a glass rod is rubbed with silk, a positive charge appears on the rod. Measurement shows that a negative charge of equal magnitude appears on the silk. This suggests that rubbing does not create charge but merely transfers it from one object to another, disturbing slightly the electrical neutrality of each. This hypothesis of the *conservation of charge* has stood up under close experimental scrutiny both for large-scale events and at the atomic and nuclear level; no exceptions have ever been found.

An interesting example of charge conservation comes about when an electron (charge $= -e$) and a positron (charge $= +e$) are brought close to each other. The two particles may simply disappear, converting all their rest mass into energy according to the well-known $E = mc^2$ relationship; this *annihilation* process was described in Section 8–9. The energy appears in the form of two oppositely directed *gamma rays*, which are similar in character to X-rays. The net charge is zero both before and after the event so that charge is conserved. Rest mass is *not* conserved, being turned completely into energy.

Another example of charge conservation is found in radioactive decay, of which the following process is typical:

$$U^{238} \rightarrow Th^{234} + He^{4}. \tag{26-5}$$

The radioactive "parent" nucleus, U^{238}, contains 92 protons (that is, its atomic number $Z = 92$). It disintegrates spontaneously by emitting an α-particle (He^4; $Z = 2$) transmuting itself into the nucleus Th^{234}, with $Z = 90$. Thus the amount of charge present before disintegration ($+92e$) is the same as that present after the disintegration.

A final example of charge conservation is found in nuclear reactions, of which the bombardment of Ca^{44} with cyclotron-accelerated protons is typical. In a particular collision a neutron may emerge from the nucleus, leaving Sc^{44} as a "residual" nucleus:

$$Ca^{44} + p \rightarrow Sc^{44} + n.$$

The sum of the atomic numbers before the reaction $(20 + 1)$ is exactly equal to the sum of the atomic numbers after the reaction $(21 + 0)$. Again charge is conserved.

QUESTIONS

1. You are given two metal spheres mounted on portable insulating supports. Find a way to give them equal and opposite charges. You may use a glass rod rubbed with silk but may not touch it to the spheres. Do the spheres have to be of equal size for your method to work?

2. A charged rod attracts bits of dry cork dust which, after touching the rod, often jump violently away from it. Explain.

3. If a charged glass rod is held near one end of an insulated uncharged metal rod as in Fig. 26–6, electrons are drawn to one end, as shown. Why does the flow of electrons cease? There is an almost inexhaustible supply of them in the metal rod.

4. In Fig. 26-6 does any net electrical force act on the metal rod? Explain.

5. An insulated rod is said to carry an electric charge. How could you verify this and determine the sign of the charge?

6. Why do electrostatic experiments not work well on humid days?

7. A person standing on an insulated stool touches a charged, insulated conductor. Is the conductor discharged completely?

8. (*a*) A positively charged glass rod attracts a suspended object. Can we conclude that the object is negatively charged? (*b*) A positively charged glass rod *repels* a suspended object. Can we conclude that the object is *positively* charged?

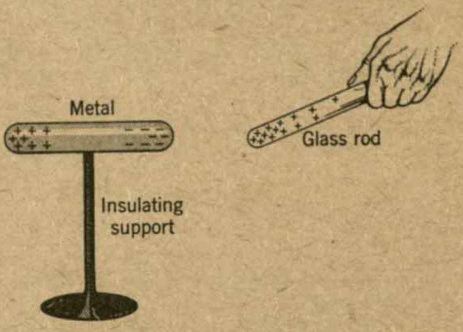


Fig. 26-6

9. Is the Coulomb force that one charge exerts on another changed if other charges are brought nearby?

10. The quantum of charge is 1.60×10^{-19} coul. Is there a corresponding single quantum of mass?

11. Verify the fact that the decay schemes for the elementary particles in Appendix E are consistent with charge conservation.

12. What does it mean to say that a physical quantity is (*a*) quantized or (*b*) conserved? Give some examples.

13. A nucleus U^{238} splits into two identical parts. Are the nuclei so produced likely to be stable or radioactive?

PROBLEMS

1. Protons in the cosmic rays strike the earth's upper atmosphere at a rate, averaged over the earth's surface, of 0.15 protons/cm^2-sec. What total current does the earth receive from beyond its atmosphere in the form of incident cosmic ray protons? The earth's radius is 6.4×10^6 meters.

2. A point charge of $+3.0 \times 10^{-6}$ coul is 12 cm distant from a second point charge of -1.5×10^{-6} coul. Calculate the magnitude and direction of the force on each charge.

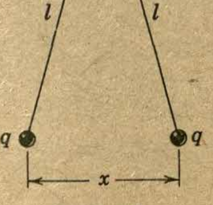


Fig. 26-7

3. Two similar balls of mass m are hung from silk threads of length l and carry similar charges q as in Fig. 26-7. Assume that θ is so small that $\tan\theta$ can be replaced by its approximate equal, $\sin\theta$. To this approximation show that

$$x = \left(\frac{q^2 l}{2\pi\epsilon_0 mg}\right)^{1/3}$$

where x is the separation between the balls. If $l = 120$ cm, $m = 10$ gm, and $x = 5.0$ cm, what is q?

4. Assume that each ball in Problem 3 is losing charge at the rate of 1.0×10^{-9} coul/sec. At what instantaneous relative speed ($= dx/dt$) do the balls approach each other initially?

5. Three small balls, each of mass 10 gm, are suspended separately from a common point by silk threads, each 1.0 meter

long. The balls are identically charged and hang at the corners of an equilateral triangle 0.1 meter long on a side. What is the charge on each ball?

6. In Fig. 26–8 what is the resultant force on the charge in the lower left corner of the square? Assume that $q = 1.0 \times 10^{-7}$ coul and $a = 5.0$ cm.

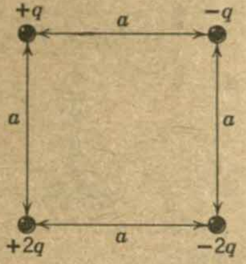


Fig. 26–8

7. A charge Q is placed at each of two opposite corners of a square. A charge q is placed at each of the other two corners. (*a*) If the resultant electrical force on Q is zero, how are Q and q related? (*b*) Could q be chosen to make the resultant force on *every* charge zero?

8. How far apart must two protons be if the electrical repulsive force acting on either one is equal to its weight?

9. (*a*) What equal positive charges would have to be placed on the earth and on the moon to neutralize their gravitational attraction? (*b*) Do you need to know the lunar distance to solve this problem? (*c*) How many pounds of hydrogen would be needed to provide the positive charge calculated in *a*?

10. A certain charge Q is to be divided into two parts, q and $Q - q$. What is the relationship of Q to q if the two parts, placed a given distance apart, are to have a maximum Coulomb repulsion?

11. Each of two small spheres is charged positively, the combined charge being 5.0×10^{-5} coul. If each sphere is repelled from the other by a force of 1.0 newton when the spheres are 2.0 meters apart, how is the total charge distributed between the spheres?

12. Two equal positive point charges are separated by a distance $2a$. A point test charge is located in a plane which is normal to the line joining these charges and midway between them. (*a*) Calculate the radius r of the circle of symmetry in this plane for which the force on the test charge has a maximum value. (*b*) What is the direction of this force, assuming a positive test charge.

13. A cube of edge a carries a point charge q at each corner. (*a*) Show that the magnitude of the resultant force on any one of the charges is

$$F = \frac{0.261q^2}{\epsilon_0 a^2}.$$

(*b*) What is the direction of $\mathbf{F}$ relative to the cube edges?

14. Estimate roughly the number of coulombs of positive charge in a glass of water.

15. (*a*) How many electrons would have to be removed from a penny to leave it with a charge of $+10^{-7}$ coul? (*b*) What fraction of the electrons in the penny does this correspond to?

16. The radius of a copper nucleus is about 1.9×10^{-13} cm. Calculate the density of the material that makes up the nucleus. Does your answer seem reasonable? (The atomic weight of copper is 64 gm/mole; ignore the mass of the electrons in comparison to that of the nucleus.)

17. In the radioactive decay of U^{238} (see Eq. 26–5) the center of the emerging α-particle is, at a certain instant, 9×10^{-15} meter from the center of the residual nucleus Th^{234}. At this instant (*a*) what is the force on the α-particle and (*b*) what is its acceleration?

The Electric Field

CHAPTER 27

27–1 The Electric Field

With every point in space near the earth we can associate a *gravitational field strength* vector $\mathbf{g}$ (see Eq. 16–12). This is the gravitational acceleration that a test body, placed at that point and released, would experience. If m is the mass of the body and $\mathbf{F}$ the gravitational force acting on it, $\mathbf{g}$ is given by

$$\mathbf{g} = \mathbf{F}/m. \tag{27–1}$$

This is an example of a *vector field*. For points near the surface of the earth the field is often taken as *uniform;* that is, $\mathbf{g}$ is the same for all points.

The flow of water in a river provides another example of a vector field, called a *flow field* (see Section 18–7). Every point in the water has associated with it a vector quantity, the velocity $\mathbf{v}$ with which the water flows past the point. If $\mathbf{g}$ and $\mathbf{v}$ do not change with time, the corresponding fields are described as *stationary*. In the case of the river note that even though the water is moving the vector $\mathbf{v}$ at any point does not change with time for steady-flow conditions.

The space surrounding a charged rod is affected by the presence of the rod, and we speak of an *electric field* in this space. In the same way we speak of a *magnetic field* in the space around a bar magnet. In the classical theory of electromagnetism the electric and magnetic fields are central concepts.

Before Faraday's time, the force acting between charged particles was thought of as a direct and instantaneous interaction between the two particles. This *action-at-a-distance* view was also held for magnetic and for gravitational forces. Today we prefer to think in terms of electric fields as follows:

1. Charge q_1 in Fig. 27–1 sets up an electric field in the space around itself. This field is suggested by the shading in the figure; later we shall show how to represent electric fields more concretely.

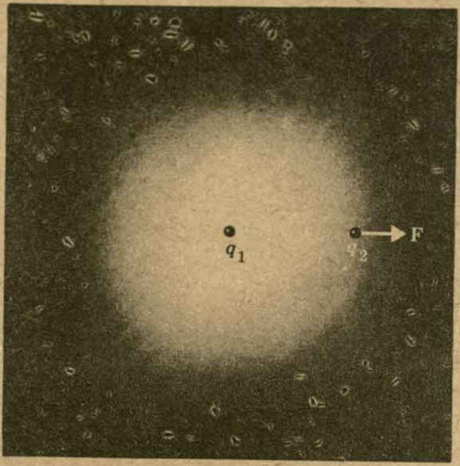


Fig. 27–1 Charge q_1 sets up a field that exerts a force **F** on charge q_2.

2. The field acts on charge q_2; this shows up in the force **F** that q_2 experiences.

The field plays an intermediary role in the forces between charges. There are two separate problems: (*a*) calculating the fields that are set up by given distributions of charge and (*b*) calculating the forces that given fields will exert on charges placed in them. We think in terms of

$$\text{charge} \rightleftharpoons \text{field}$$

and not, as in the action-at-a-distance point of view, in terms of

$$\text{charge} \rightleftharpoons \text{charge.}$$

In Fig. 27–1 we can also imagine that q_2 sets up a field and that this field acts on q_1, producing a force $-\mathbf{F}$ on it. The situation is completely symmetrical, each charge being immersed in a field associated with the other charge.

If the only problem in electromagnetism was that of the forces between stationary charges, the field and the action-at-a-distance points of view would be perfectly equivalent. Suppose, however, that q_1 in Fig. 27–1 suddenly accelerates to the right. How quickly does the charge q_2 learn that q_1 has moved and that the force which it (q_2) experiences must increase? Electromagnetic theory predicts that q_2 learns about q_1's motion by a *field disturbance* that emanates from q_1, traveling with the speed of light. The action-at-a-distance point of view requires that information about q_1's acceleration be communicated *instantaneously* to q_2; this is not in accord with experiment. Accelerating electrons in the antenna of a radio transmitter influence electrons in a distant receiving antenna only after a time l/c where l is the separation of the antennas and c is the speed of light.

27-2 The Electric Field Strength E

To define the electric field operationally, we place a small test body carrying a test charge q_0 (assumed positive for convenience) at the point in space that is to be examined, and we measure the electrical force **F** (if any) that acts on this body. The *electric field strength* **E** at the point is defined as *

$$\mathbf{E} = \mathbf{F}/q_0. \tag{27-2}$$

Here **E** is a vector because **F** is one, q_0 being a scalar. The direction of **E** is the direction of **F**, that is, it is the direction in which a resting positive charge placed at the point would tend to move.

The definition of gravitational field strength **g** is much like that of electric field strength, except that the mass of the test body rather than its charge is the property of interest. Although the units of **g** are usually written as meters/sec^2, they could also be written as nt/kg (Eq. 27-1); those for **E** are nt/coul (Eq. 27-2). Thus both **g** and **E** are expressed as a force divided by a property (mass or charge) of the test body.

▶ **Example 1.** What is the magnitude of the electric field strength **E** such that an electron, placed in the field, would experience an electrical force equal to its weight?

From Eq. 27-2, replacing q_0 by e and F by mg, where m is the electron mass, we have

$$\begin{aligned} E &= \frac{F}{q_0} = \frac{mg}{e} \\ &= \frac{(9.1 \times 10^{-31}\ \text{kg})(9.8\ \text{meters/sec}^2)}{1.6 \times 10^{-19}\ \text{coul}} \\ &= 5.6 \times 10^{-11}\ \text{nt/coul}. \end{aligned}$$

This is a very weak electric field. Which way will **E** have to point if the electric force is to cancel the gravitational force? ◀

In applying Eq. 27-2 we must use a test charge as small as possible. A large test charge might disturb the primary charges that are responsible for the field, thus changing the very quantity that we are trying to measure. Equation 27-2 should, strictly, be replaced by

$$\mathbf{E} = \lim_{q_0 \to 0} \frac{\mathbf{F}}{q_0}. \tag{27-3}$$

This equation instructs us to use a smaller and smaller test charge q_0, evaluating the ratio $\mathbf{F}/q_0$ at every step. The electric field **E** is then the limit of this ratio as the size of the test charge approaches zero.

27-3 Lines of Force

The concept of the electric field as a vector was not appreciated by Michael Faraday, who always thought in terms of *lines of force*. The lines of force still form a convenient way of visualizing electric-field patterns. We shall use them for this purpose but we shall not employ them quantitatively.

---

* This definition of **E**, though conceptually sound and quite appropriate to our present purpose, is rarely carried out in practice because of experimental difficulties. **E** is normally found by calculation from more readily measurable quantities such as the electric potential; see Section 29-7.

The relationship between the (imaginary) lines of force and the electric field strength vector is this:

1. The tangent to a line of force at any point gives the *direction* of **E** at that point.
2. The lines of force are drawn so that the number of lines per unit cross-sectional area is proportional to the *magnitude* of **E**. Where the lines are close together E is large and where they are far apart E is small.

It is not obvious that it is possible to draw a continuous set of lines to meet these requirements. Indeed, it turns out that if Coulomb's law were not true it would *not* be possible to do so; see Problem 4.

Figure 27–2 shows the lines of force for a uniform sheet of positive charge. We assume that the sheet is infinitely large, which, for a sheet of finite dimensions, is equivalent to considering only those points whose distance from the sheet is small compared to the distance to the nearest edge of the sheet. A positive test charge, released in front of such a sheet, would move away from the sheet along a perpendicular line. Thus the electric field strength vector at any point near the sheet must be at right angles to the sheet. The lines of force are uniformly spaced, which means that **E** has the same magnitude for all points near the sheet.

Figure 27–3 shows the lines of force for a negatively charged sphere. From symmetry, the lines must lie along radii. They point inward because a free positive charge would be accelerated in this direction. The electric field E is not constant but decreases with increasing distance from the charge. This is evident in the lines of force, which are farther apart at greater distances. From symmetry, E is the same for all points that lie a given distance from the center of the charge.

▶ **Example 2.** In Fig. 27–3 how does E vary with the distance r from the center of the charged sphere?

Suppose that N lines originate on the sphere. Draw an imaginary concentric sphere of radius r; the number of lines per unit cross-sectional area at every point on the sphere is $N/4\pi r^2$. Since E is proportional to this, we can write that

$$E \propto 1/r^2.$$

We derive an exact relationship in Section 27–4. How should E vary with distance from an infinitely long uniform cylinder of charge? ◀

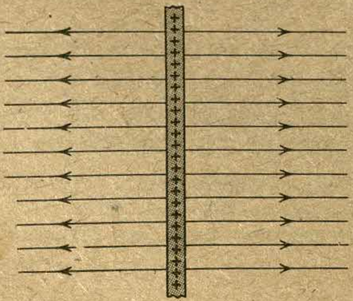


Fig. 27–2 Lines of force for a section of an infinitely large sheet of positive charge.

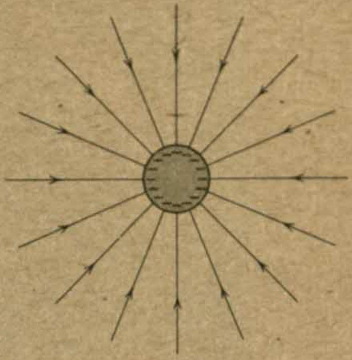


Fig. 27-3 Lines of force for a negatively charged sphere.

Figures 27-4 and 27-5 show the lines of force for two equal like charges and for two equal unlike charges, respectively. Michael Faraday, as we have said, used lines of force a great deal in his thinking. They were more real for him than they are for most scientists and engineers today. It is possible

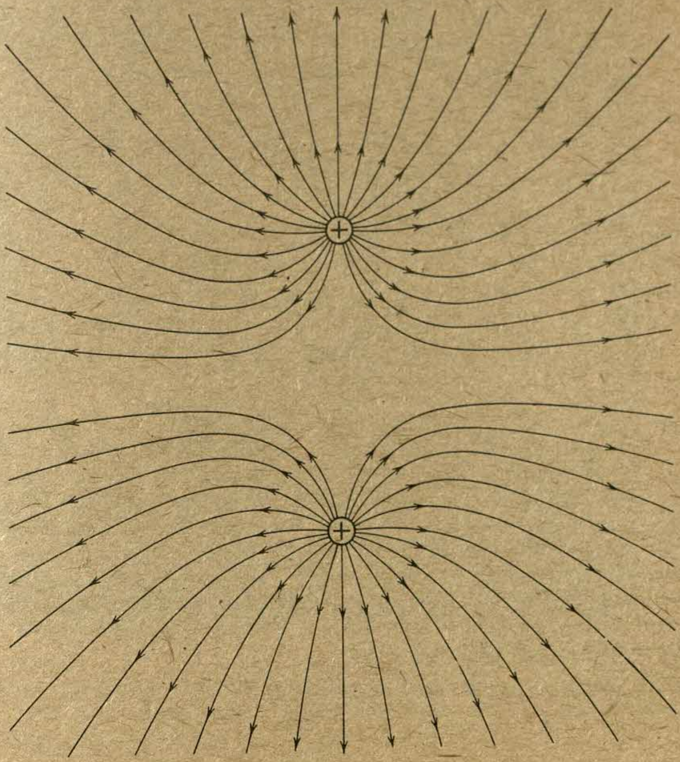


Fig. 27-4 Lines of force for two equal positive charges.

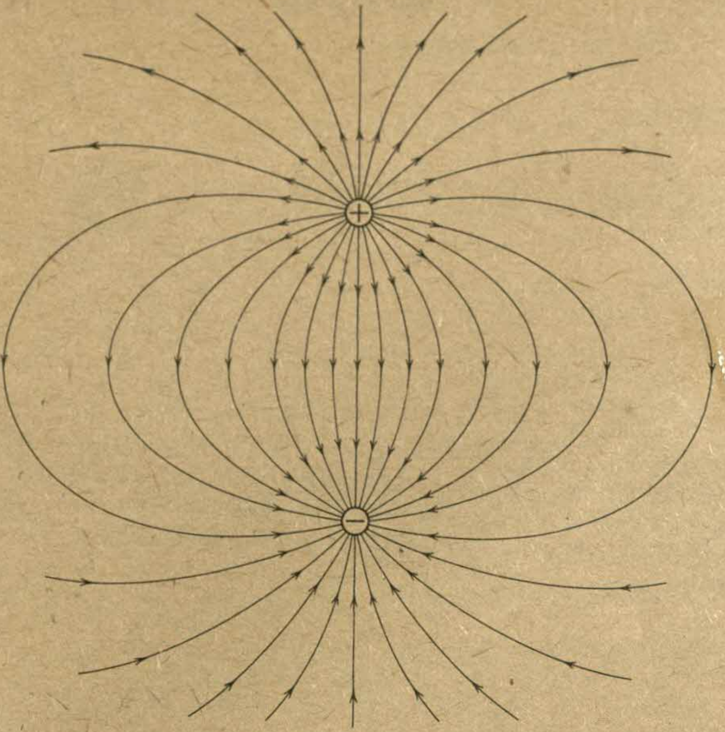


Fig. 27–5 Lines of force for equal but opposite charges.

to sympathize with Faraday's point of view. Can we not almost "see" the charges being pushed apart in Fig. 27–4 and pulled together in Fig. 27–5 by the lines of force? The student should compare Fig. 27–5 with Fig. 18–15, which represents a flow field. Figure 27–6 shows a representation of lines of force around charged conductors, using grass seeds suspended in an insulating liquid.

Lines of force give a vivid picture of the way **E** varies through a given region of space. However, the equations of electromagnetism (see Table 38–3) are written in terms of the electric field strength **E** and other field vectors and not in terms of the lines of force. The electric field **E** varies in a perfectly continuous way as any path in the field is traversed; see Fig. 27-7*a*. This kind of physical continuity, however, is different from the continuity of the lines of force, which has no real physical meaning. To illustrate this point we assert that the lines of force need not be continuous but can be drawn as in Fig. 27–7*b*. This figure agrees with the two properties of lines of force laid down at the beginning of this section. This point of view has been especially emphasized by Joseph Slepian.*

---

* See *American Journal of Physics*, **19**, **87** (1951).

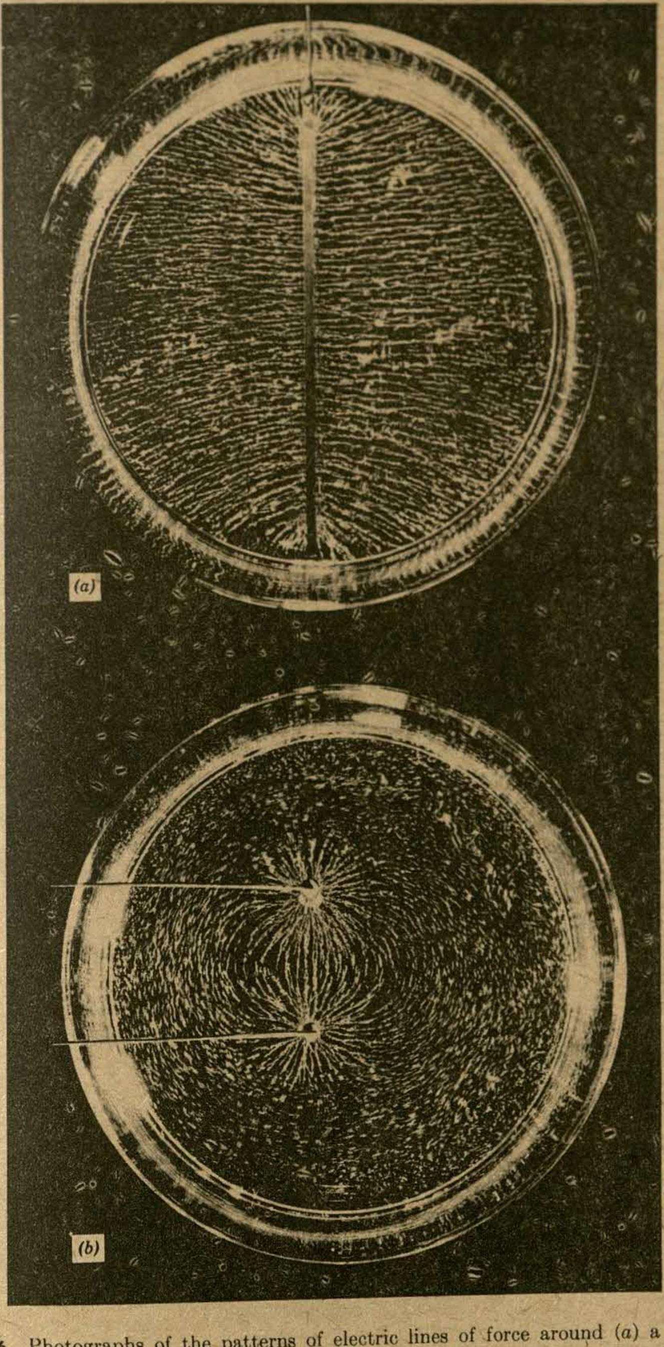
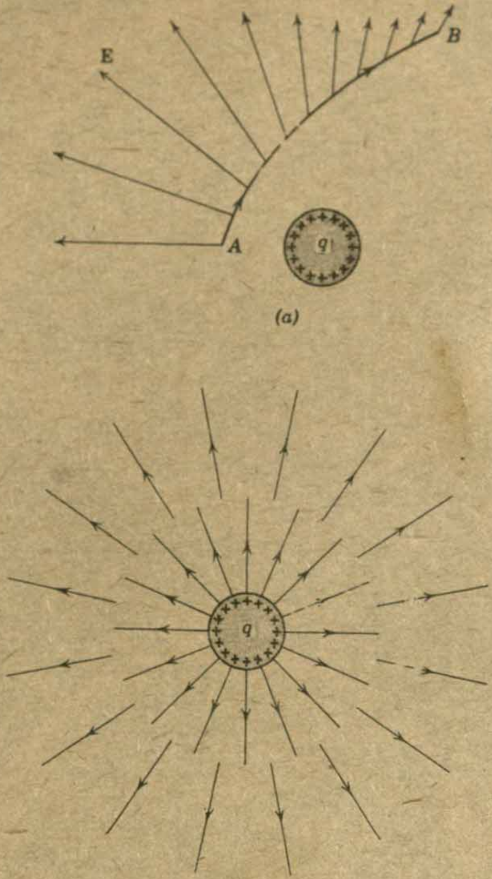


Fig. 27–6 Photographs of the patterns of electric lines of force around (*a*) a charged plate (compare Fig. 27–2), and (*b*) two rods with equal and opposite charges (compare Fig. 27–5). The patterns were made by suspending grass seed in an insulating liquid. (Courtesy Educational Services Incorporated, Watertown, Mass.)



(a)

(b)

Fig. 27-7 (*a*) **E** varies continuously as we move along any path *AB* in the field set up by *q*. (*b*) Lines of force need not be continuous.

27-4 Calculation of E

Let a test charge q_0 be placed a distance r from a point charge q. The magnitude of the force acting on q_0 is given by Coulomb's law, or

$$F = \frac{1}{4\pi\epsilon_0}\frac{qq_0}{r^2}.$$

The electric field strength at the site of the test charge is given by Eq. 27-2, or

$$E = \frac{F}{q_0} = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2}. \qquad (27\text{-}4)$$

The direction of **E** is on a radial line from q, pointing outward if q is positive and inward if q is negative.

To find **E** for a group of point charges: (*a*) Calculate $\mathbf{E}_n$ due to each charge at the given point *as if it were the only charge present*. (*b*) Add these separately calculated fields vectorially to find the resultant field **E** at the point. In equation form,

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 + \mathbf{E}_3 + \cdots = \Sigma \mathbf{E}_n \qquad n = 1, 2, 3, \ldots. \tag{27-5}$$

The sum is a vector sum, taken over all the charges.

If the charge distribution is a continuous one, the field it sets up at any point P can be computed by dividing the charge into infinitesimal elements dq. The field $d\mathbf{E}$ due to each element at the point in question is then calculated, treating the elements as point charges. The magnitude of $d\mathbf{E}$ (see Eq. 27–4) is given by

$$dE = \frac{1}{4\pi\epsilon_0}\frac{dq}{r^2}, \tag{27-6}$$

where r is the distance from the charge element dq to the point P. The resultant field at P is then found by adding (that is, integrating) the field contributions due to all the charge elements, or,

$$\mathbf{E} = \int d\mathbf{E}. \tag{27-7}$$

The integration, like the sum in Eq. 27–5, is a vector operation; in Example 5 we will see how such an integral is handled in a simple case.

▶ **Example 3.** *An electric dipole.* Figure 27–8 shows a positive and a negative charge of equal magnitude q placed a distance $2a$ apart, a configuration called an electric dipole. The pattern of lines of force is that of Fig. 27–5, which also shows an electric dipole. What is the field **E** due to these charges at point P, a distance r along the perpendicular bisector of the line joining the charges? Assume $r \gg a$.

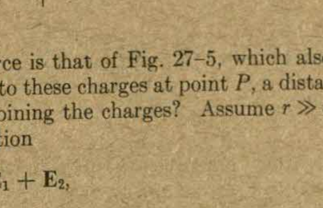


Fig. 27–8 Example 3.

Equation 27–5 gives the vector equation

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2,$$

where, from Eq. 27–4,*

$$E_1 = E_2 = \frac{1}{4\pi\epsilon_0}\frac{q}{a^2 + r^2}.$$

---

* Note that the r's in Eq. 27–4 and in this equation have different meanings.

16335

The vector sum of $\mathbf{E}_1$ and $\mathbf{E}_2$ points vertically downward and has the magnitude

$$E = 2E_1 \cos\theta.$$

From the figure we see that

$$\cos\theta = \frac{a}{\sqrt{a^2 + r^2}}.$$

Substituting the expressions for E_1 and for $\cos\theta$ into that for E yields

$$E = \frac{2}{4\pi\epsilon_0}\frac{q}{(a^2 + r^2)}\frac{a}{\sqrt{a^2 + r^2}} = \frac{1}{4\pi\epsilon_0}\frac{2aq}{(a^2 + r^2)^{3/2}}.$$

If $r \gg a$, we can neglect a in the denominator; this equation then reduces to

$$E \cong \frac{1}{4\pi\epsilon_0}\frac{(2a)(q)}{r^3}. \tag{27-8a}$$

The essential properties of the charge distribution in Fig. 27–8, the magnitude of the charge q and the separation $2a$ between the charges, enter Eq. 27–8*a* only as a product. This means that, if we measure $\mathbf{E}$ at various distances from the electric dipole (assuming $r \gg a$), we can never deduce q and $2a$ separately but only the product $2aq$; if q were doubled and a simultaneously cut in half, the electric field *at large distances from the dipole* would not change.

The product $2aq$ is called the *electric dipole moment* p. Thus we can rewrite this equation for E, *for distant points along the perpendicular bisector*, as

$$E = \frac{1}{4\pi\epsilon_0}\frac{p}{r^3}. \tag{27-8b}$$

The result for distant points *along the dipole axis* (see Problem 10) and the general result for any distant point (see Problem 23) also contain the quantities $2a$ and q only as the product $2aq$ ($= p$). The variation of E with r in the general result for distant points is also as $1/r^3$, as in Eq. 27–8*b*.

The dipole of Fig. 27–8 is two equal and opposite charges placed close to each other so that their separate fields at distant points almost, but not quite, cancel. On this point of view it is easy to understand that $E(r)$ for a dipole varies as $1/r^3$ (Eq. 27–8*b*), whereas for a point charge $E(r)$ drops off more slowly, namely as $1/r^2$ (Eq. 27–4).

Example 4. Figure 27–9 shows a charge q_1 ($= +1.0 \times 10^{-6}$ coul) 10 cm from a charge q_2 ($= +2.0 \times 10^{-6}$ coul). At what point on the line joining the two charges is the electric field strength zero?

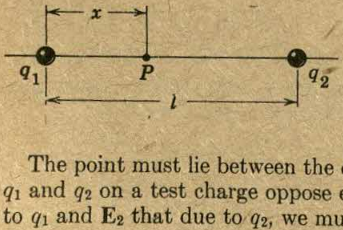


Fig. 27–9 Example 4.

The point must lie between the charges because only here do the forces exerted by q_1 and q_2 on a test charge oppose each other. If $\mathbf{E}_1$ is the electric field strength due to q_1 and $\mathbf{E}_2$ that due to q_2, we must have

$$E_1 = E_2$$

or (see Eq. 27–4)

$$\frac{1}{4\pi\epsilon_0}\frac{q_1}{x^2} = \frac{1}{4\pi\epsilon_0}\frac{q_2}{(l - x)^2}$$

where x is the distance from q_1 and l equals 10 cm. Solving for x yields

$$x = \frac{l}{1 + \sqrt{q_2/q_1}} = \frac{10 \text{ cm}}{1 + \sqrt{2}} = 4.1 \text{ cm}.$$

The student should supply the missing steps. On what basis was the second root of the resulting quadratic equation discarded?

Example 5. Figure 27–10 shows a ring of charge q and of radius a. Calculate **E** for points on the axis of the ring a distance x from its center.

Consider a differential element of the ring of length ds, located at the top of the ring in Fig. 27–10. It contains an element of charge given by

$$dq = q\frac{ds}{2\pi a}$$

where $2\pi a$ is the circumference of the ring. This element sets up a differential electric field $d\mathbf{E}$ at point P.

The resultant field **E** at P is found by integrating the effects of all the elements that make up the ring. From symmetry this resultant field must lie along the ring axis. Thus only the component of $d\mathbf{E}$ parallel to this axis contributes to the final result. The component perpendicular to the axis is canceled out by an equal but opposite component established by the charge element on the opposite side of the ring.

Thus the general vector integral (Eq. 27–7)

$$\mathbf{E} = \int d\mathbf{E}$$

becomes a scalar integral $$E = \int dE \cos\theta.$$

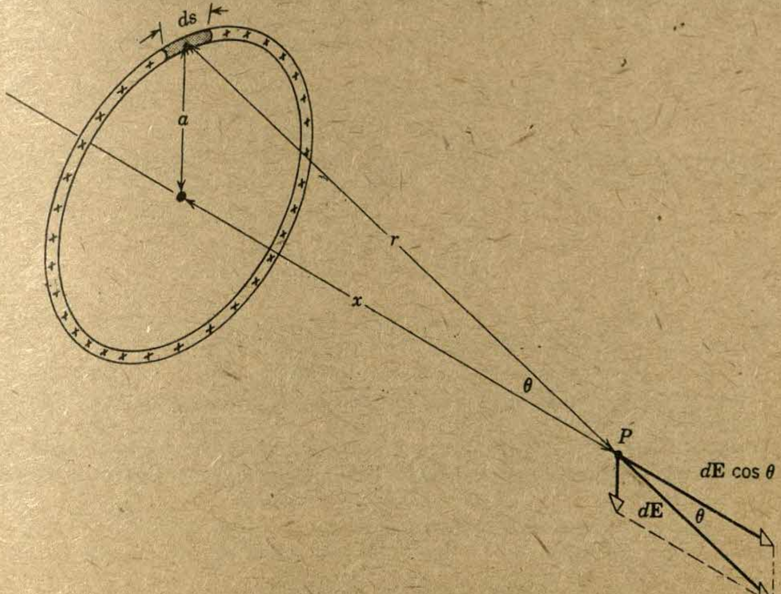


Fig. 27–10 Example 5

The quantity dE follows from Eq. 27-6, or

$$dE = \frac{1}{4\pi\epsilon_0}\frac{dq}{r^2} = \frac{1}{4\pi\epsilon_0}\left(\frac{q\,ds}{2\pi a}\right)\frac{1}{a^2+x^2}.$$

From Fig. 27-10 we have $$\cos\theta = \frac{x}{\sqrt{a^2+x^2}}.$$

Noting that, for a given point P, x has the same value for all charge elements and is not a variable, we obtain

$$E = \int dE\cos\theta = \int \frac{1}{4\pi\epsilon_0}\frac{q\,ds}{(2\pi a)(a^2+x^2)}\frac{x}{\sqrt{a^2+x^2}}$$

$$= \frac{1}{4\pi\epsilon_0}\frac{qx}{(2\pi a)(a^2+x^2)^{3/2}}\int ds.$$

The integral is simply the circumference of the ring ($= 2\pi a$), so that

$$E = \frac{1}{4\pi\epsilon_0}\frac{qx}{(a^2+x^2)^{3/2}}.$$

Does this expression for E reduce to an expected result for $x = 0$? For $x \gg a$ we can neglect a in the denominator of this equation, yielding

$$E \cong \frac{1}{4\pi\epsilon_0}\frac{q}{x^2}.$$

This is an expected result (compare Eq. 27-4) because at great enough distances the ring behaves like a point charge q.

Example 6. *Line of charge.* Figure 27-11 shows a section of an infinite line of charge whose linear charge density (that is, the charge per unit length, measured in coul/meter) has the constant value λ. Calculate the field $\mathbf{E}$ a distance y from the line.

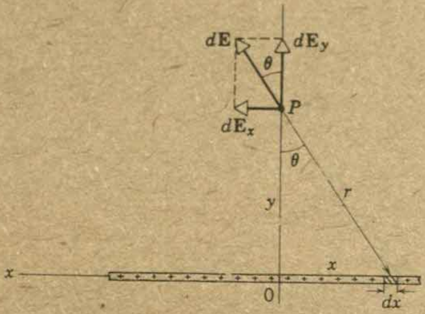


Fig. 27-11 Example 6. A section of an infinite line of charge.

The magnitude of the field contribution dE due to charge element dq ($= \lambda\,dx$) is given, using Eq. 27-6, by

$$dE = \frac{1}{4\pi\epsilon_0}\frac{dq}{r^2} = \frac{1}{4\pi\epsilon_0}\frac{\lambda\,dx}{y^2+x^2}.$$

The vector $d\mathbf{E}$, as Fig. 27-11 shows, has the components

$$dE_x = -dE\sin\theta \quad \text{and} \quad dE_y = dE\cos\theta.$$

The minus sign in front of dE_x indicates that $d\mathbf{E}_x$ points in the negative x direction. The x and y components of the resultant vector $\mathbf{E}$ at point P are given by

$$E_x = \int dE_x = -\int_{x=-\infty}^{x=+\infty} \sin\theta\, dE \quad \text{and} \quad E_y = \int dE_y = \int_{x=-\infty}^{x=+\infty} \cos\theta\, dE.$$

E_x must be zero because every charge element on the right has a corresponding element on the left such that their field contributions in the x direction cancel. Thus $\mathbf{E}$ points entirely in the y direction. Because the contributions to E_y from the right- and left-hand halves of the rod are equal, we can write

$$E = E_y = 2\int_{x=0}^{x=+\infty} \cos\theta\, dE.$$

Note that we have changed the lower limit of integration and have introduced a factor of two.

Substituting the expression for dE into this equation gives

$$E = \frac{\lambda}{2\pi\epsilon_0}\int_{x=0}^{x=\infty} \cos\theta\, \frac{dx}{y^2 + x^2}.$$

From Fig. 27-11 we see that the quantities θ and x are not independent. We must eliminate one of them, say x. The relation between x and θ is (see figure)

$$x = y\tan\theta.$$

Differentiating, we obtain $$dx = y\sec^2\theta\, d\theta.$$

Substituting these two expressions leads finally to

$$E = \frac{\lambda}{2\pi\epsilon_0 y}\int_{\theta=0}^{\theta=\pi/2} \cos\theta\, d\theta.$$

The student should check this step carefully, noting that the limits must now be on θ and not on x. For example as $x \to +\infty$, $\theta \to \pi/2$, as Fig. 27-11 shows. This equation integrates readily to

$$E = \frac{\lambda}{2\pi\epsilon_0 y}\Big|\sin\theta\Big|_0^{\pi/2} = \frac{\lambda}{2\pi\epsilon_0 y}.$$

The student may wonder about the usefulness of solving a problem involving an infinite rod of charge when any actual rod must have a finite length (see Problem 15). However, for points close enough to finite rods and not near their ends, the equation that we have just derived yields results that are so close to the correct values that the difference can be ignored in many practical situations. It is usually unnecessary to solve exactly every geometry encountered in practical problems. Indeed, if idealizations or approximations are not made, the vast majority of significant problems of all kinds in physics and engineering cannot be solved at all. ◀

27-5 A Point Charge in an Electric Field

An electric field will exert a force on a charged particle given by (Eq. 27-2)

$$\mathbf{F} = \mathbf{E}q.$$

This force will produce an acceleration

$$\mathbf{a} = \mathbf{F}/m,$$

where m is the mass of the particle. We will consider two examples of the acceleration of a charged particle in a uniform electric field. Such a field can be produced by connecting the terminals of a battery to two parallel metal plates which are otherwise insulated from each other. If the spacing between the plates is small compared with the dimensions of the plates, the field between them will be fairly uniform except near the edges. Note that in calculating the motion of a particle in a field set up by external charges the field due to the particle itself (that is, its *self-field*) is ignored. For example, the earth's gravitational field can have no effect on the earth itself but only on a second object, say a stone, placed in that field.

▶ **Example 7.** A particle of mass m and charge q is placed at rest in a uniform electric field (Fig. 27–12) and released. Describe its motion.

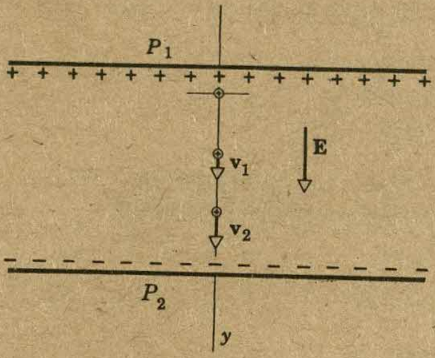


Fig. 27–12 A charge is released from rest in a uniform electric field set up between two oppositely charged metal plates P_1 and P_2.

The motion resembles that of a body falling in the earth's gravitational field. The (constant) acceleration is given by

$$a = \frac{F}{m} = \frac{qE}{m}.$$

The equations for uniformly accelerated motion (Table 3–1) then apply. With $v_0 = 0$, they are

$$v = at = \frac{qEt}{m},$$

$$y = \tfrac{1}{2}at^2 = \frac{qEt^2}{2m},$$

and

$$v^2 = 2ay = \frac{2qEy}{m}.$$

The kinetic energy attained after moving a distance y is found from

$$K = \tfrac{1}{2}mv^2 = \tfrac{1}{2}m\left(\frac{2qEy}{m}\right) = qEy.$$

This result also follows directly from the work-energy theorem because a constant force qE acts over a distance y.

Example 8. *Deflecting an electron beam.* Figure 27–13 shows an electron of mass m and charge e projected with speed v_0 at right angles to a uniform field **E**. Describe its motion.

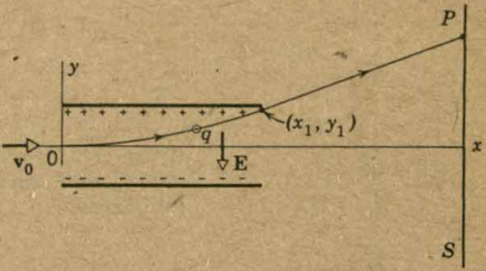


Fig. 27–13 Example 8. An electron is projected into a uniform electric field.

The motion is like that of a projectile fired horizontally in the earth's gravitational field. The considerations of Section 4–3 apply, the horizontal (x) and vertical (y) motions being given by

$$x = v_0t$$

and

$$y = \tfrac{1}{2}at^2 = \frac{eE}{2m}t^2.$$

Eliminating t yields

$$y = \frac{eE}{2mv_0{}^2}x^2 \tag{27–9}$$

for the equation of the trajectory.

When the electron emerges from the plates in Fig. 27–13, it travels (neglecting gravity) in a straight line tangent to the parabola at the exit point. We can let it fall on a fluorescent screen S placed some distance beyond the plates. Together with other electrons following the same path, it will then make itself visible as a small luminous spot; this is the principle of the electrostatic *cathode-ray oscilloscope.*

Example 9. The electric field between the plates of a cathode-ray oscilloscope is 1.2×10^4 nt/coul. What deflection will an electron experience if it enters at right angles to the field with a kinetic energy of 2000 ev ($= 3.2 \times 10^{-16}$ joule), a typical value? The deflecting assembly is 1.5 cm long.

Recalling that $K_0 = \frac{1}{2}mv_0{}^2$, we can rewrite Eq. 27–9 as

$$y = \frac{eEx^2}{4K_0}.$$

If x_1 is the horizontal position of the far edge of the plate, y_1 will be the corresponding deflection (see Fig. 27–13), or

$$\begin{aligned} y_1 &= \frac{eEx_1{}^2}{4K_0} \\ &= \frac{(1.6 \times 10^{-19}\text{ coul})(1.2 \times 10^4\text{ nt/coul})(1.5 \times 10^{-2}\text{ meter})^2}{(4)(3.2 \times 10^{-16}\text{ joule})} \\ &= 3.4 \times 10^{-4}\text{ meter} = 0.34\text{ mm}. \end{aligned}$$

The deflection measured, not at the deflecting plates but at the fluorescent screen, is much larger.

Example 10. A positive point test charge q_0 is placed halfway between two equa itive charges q. What force acts on it at or near this point P?

rom symmetry the force *at* the point is zero so that the particle is in equilibrium; nature of the equilibrium remains to be found. Figure 27–14 (compare Fig. 27–4)

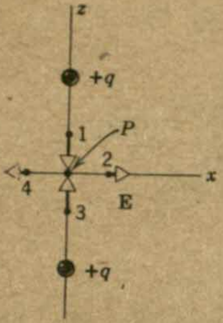


Fig. 27–14 Example 10. The electric field at four points near a point P which is centered between two equal positive charges.

shows the **E** vectors for four points near P. If the test charge is moved along the z axis, a *restoring* force is brought into play; however, the equilibrium is unstable for motion in the x-y plane. Thus we have the three-dimensional equivalent of *saddle point equilibrium;* see Fig. 14–8. What is the nature of the equilibrium for a negative test charge? ◀

27–6 A Dipole in an Electric Field

An electric dipole moment can be regarded as a vector **p** whose magnitude p, for a dipole like that described in Example 3, is the product $2aq$ of the magnitude of either charge q and the distance $2a$ between the charges. The direction of **p** for such a dipole is from the negative to the positive charge. The vector nature of the electric dipole moment permits us to cast many expressions involving electric dipoles into concise form, as we shall see.

Figure 27–15*a* shows an electric dipole formed by placing two charges $+q$ and $-q$, a fixed distance $2a$ apart. The arrangement is placed in a uniform *external* electric field **E**, its dipole moment **p** making an angle θ with this field. Two equal and opposite forces **F** and $-$**F** act as shown, where

$$F = qE.$$

The net force is clearly zero, but there is a net torque about an axis through O (see Eq. 12–2) given by

$$\tau = 2F(a \sin \theta) = 2aF \sin \theta.$$

Combining these two equations and recalling that $p = (2a)(q)$, we obtain

$$\tau = 2aqE \sin \theta = pE \sin \theta. \tag{27-10}$$

Thus an electric dipole placed in an external electric field **E** experiences a torque tending to align it with the field. Equation 27–10 can be written in vector form as

$$\boldsymbol{\tau} = \mathbf{p} \times \mathbf{E}, \tag{27-11}$$

the appropriate vectors being shown in Fig. 27–15*b*

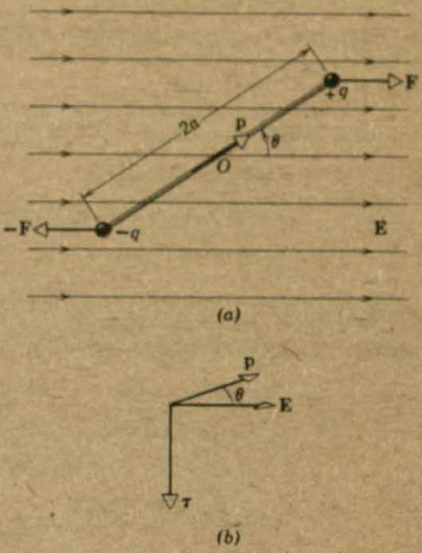


Fig. 27–15 (*a*) An electric dipole in a uniform external field. (*b*) An oblique view, illustrating $\boldsymbol{\tau} = \mathbf{p} \times \mathbf{E}$.

Work (positive or negative) must be done by an external agent to change the orientation of an electric dipole in an external field. This work is stored as potential energy U in the system consisting of the dipole and the arrangement used to set up the external field. If θ in Fig. 27–15*a* has the initial value θ_0, the work required to turn the dipole axis to an angle θ is given (see Table 12–2) from

$$W = \int dW = \int_{\theta_0}^{\theta} \tau \, d\theta = U,$$

where τ is the torque exerted by the agent that does the work. Combining this equation with Eq. 27–10 yields

$$U = \int_{\theta_0}^{\theta} pE \sin\theta \, d\theta = pE \int_{\theta_0}^{\theta} \sin\theta \, d\theta$$

$$= pE \Big| -\cos\theta \Big|_{\theta_0}^{\theta} .$$

Since we are interested only in *changes* in potential energy, we choose the reference orientation θ_0 to have any convenient value, in this case 90°. This gives

$$U = -pE \cos\theta \tag{27–12}$$

or, in vector form,

$$U = -\mathbf{p} \cdot \mathbf{E}. \tag{27–13}$$

▶ **Example 11.** An electric dipole consists of two opposite charges of magnitude $q = 1.0 \times 10^{-6}$ coul separated by $d = 2.0$ cm. The dipole is placed in an external field of 1.0×10^5 nt/coul.

(*a*) What maximum torque does the field exert on the dipole? The maximum torque is found by putting $\theta = 90°$ in Eq. 27–10 or

$$\begin{aligned}\tau &= pE\sin\theta = qdE\sin\theta \\ &= (1.0 \times 10^{-6}\text{ coul})(0.020\text{ meter})(1.0 \times 10^{5}\text{ nt/coul})(\sin 90°) \\ &= 2.0 \times 10^{-3}\text{ nt-m.}\end{aligned}$$

(*b*) How much work must an external agent do to turn the dipole end for end, starting from a position of alignment ($\theta = 0$)? The work is the difference in potential energy U between the positions $\theta = 180°$ and $\theta = 0$. From Eq. 27–12,

$$\begin{aligned}W &= U_{180°} - U_{0°} = (-pE\cos 180°) - (-pE\cos 0) \\ &= 2pE = 2qdE \\ &= (2)(1.0 \times 10^{-6}\text{ coul})(0.020\text{ meter})(1.0 \times 10^{5}\text{ nt/coul}) \\ &= 4.0 \times 10^{-3}\text{ joule.}\end{aligned}$$ ◀

QUESTIONS

1. Name as many scalar fields and vector fields as you can.

2. (*a*) In the gravitational attraction between the earth and a stone, can we say that the earth lies in the gravitational field of the stone? (*b*) How is the gravitational field due to the stone related to that due to the earth?

3. A positively charged ball hangs from a long silk thread. We wish to measure **E** at a point in the same horizontal plane as that of the hanging charge. To do so, we put a positive test charge q_0 at the point and measure $\mathbf{F}/q_0$. Will F/q_0 be less than, equal to, or greater than E at the point in question?

4. Taking into account the quantization of electric charge (the single electron providing the basic charge unit), how can we justify the procedure suggested by Eq. 27–3?

5. In Fig. 27–5 the force on the lower charge points up and is finite. The crowding of the lines of force, however, suggests that E is infinitely great at the site of this (point) charge. A charge immersed in an infinitely great field should have an infinitely great force acting on it. What is the solution to this dilemma?

6. Electric lines of force never cross. Why?

7. In Fig. 27–4 why do the lines of force around the edge of the figure appear, when extended backwards, to radiate uniformly from the center of the figure?

8. Figure 27–2 shows that **E** has the same value for all points in front of an infinite uniformly charged sheet. Is this reasonable? One might think that the field should be stronger near the sheet because the charges are so much closer.

9. Two point charges of unknown magnitude and sign are a distance d apart. The electric field strength is zero at one point between them, on the line joining them. What can you conclude about the charges?

10. Compare the way E varies with r for (*a*) a point charge (Eq. 27–4), (*b*) a dipole (Eq. 27–8*a*), and (*c*) a quadrupole (Problem 18).

11. If a point charge q of mass m is released from rest in a nonuniform field, will it follow a line of force?

12. An electric dipole is placed in a *nonuniform* electric field. Is there a net force on it?

13. An electric dipole is placed at rest in a uniform external electric field, as in Fig. 27–15*a*, and released. Discuss its motion.

14. An electric dipole has its dipole moment **p** aligned with a uniform external electric field **E**. (*a*) Is the equilibrium stable or unstable? (*b*) Discuss the nature of the equilibrium if **p** and **E** point in opposite directions.

PROBLEMS

1. Sketch qualitatively the lines of force associated with a thin, circular, uniformly charged disk of radius R. (Hint: Consider as limiting cases points very close to the surface and points very far from it.) Show the lines only in a plane containing the axis of the disk.

2. (*a*) Sketch qualitatively the lines of force associated with three equal positive point charges placed at the corners of an equilateral triangle. Use symmetry arguments and limiting cases (see hint in Problem 1). Show the lines in the plane of the triangle only. (*b*) Discuss the nature of the equilibrium of a test charge placed at the center of the triangle.

3. In Fig. 27–4 consider any two lines of force leaving the upper charge. If the angle between their tangents for points near the charge is θ, it becomes $\theta/2$ at great distances. Verify this statement and explain it. (Hint: Consider how the lines must behave both close to either charge and far from the charges.)

4. Assume that the exponent in Coulomb's law is not "two" but n. Show that for $n \neq 2$ it is impossible to construct lines that will have the properties listed for lines of force in Section 27–3. For simplicity, treat an isolated point charge.

5. What is the magnitude of a point charge chosen so that the electric field 50 cm away has the magnitude 2.0 nt/coul?

6. Two equal and opposite charges of magnitude 2.0×10^{-7} coul are 15 cm apart. (*a*) What are the magnitude and direction of **E** at a point midway between the charges? (*b*) What force (magnitude and direction) would act on an electron placed there?

7. Two point charges are a distance d apart (Fig. 27–16). Plot $E(x)$, assuming $x = 0$ at the left-hand charge. Consider both positive and negative values of x. Plot E as positive if **E** points to the right and negative if **E** points to the left. Assume $q_1 = +1.0 \times 10^{-6}$ coul, $q_2 = +3.0 \times 10^{-6}$ coul, and $d = 10$ cm.

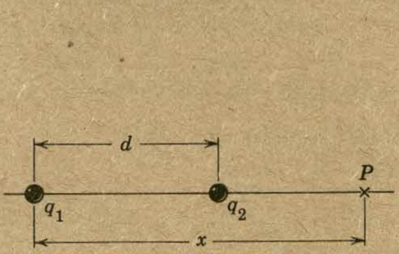


Fig. 27–16

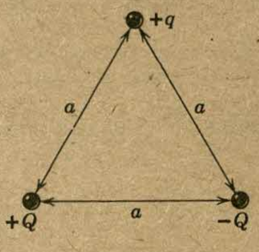


Fig. 27–17

8. Three charges are arranged in an equilateral triangle as in Fig. 27–17. What is the direction of the force on $+q$?

9. Two point charges of magnitude $+2.0 \times 10^{-7}$ coul and $+8.5 \times 10^{-8}$ coul are 12 cm apart. (*a*) What electric field does each produce at the site of the other? (*b*) What force acts on each?

10. *Axial field due to an electric dipole.* In Fig. 27–8, consider a point a distance r from the center of the dipole *along its axis*. (*a*) Show that, at large values of r, the electric field is

$$E = \frac{1}{2\pi\epsilon_0}\frac{p}{r^3},$$

which is twice the value given for the conditions of Example 3. (*b*) What is the direction of **E**?

11. In Fig. 27–8 assume that both charges are positive. (*a*) Show that E at point P in that figure, assuming $r \gg a$, is given by

$$E = \frac{1}{4\pi\epsilon_0}\frac{2q}{r^2}.$$

(*b*) What is the direction of **E**? (*c*) Is it reasonable that E should vary as r^{-2} here and as r^{-3} for the dipole of Fig. 27–8?

12. (*a*) In Fig. 27–18 locate the point (or points) at which the electric field strength is zero. (*b*) Sketch qualitatively the lines of force. Take a = 50 cm.

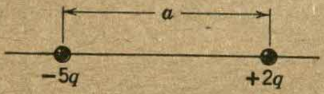


Fig. 27–18

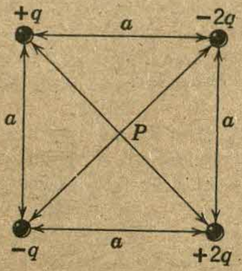


Fig. 27–19

13. What is **E** in magnitude and direction at the center of the square of Fig. 27–19? Assume that $q = 1.0 \times 10^{-8}$ coul and a = 5.0 cm.

14. Two point charges of unknown magnitude and sign are placed a distance d apart. (*a*) If it is possible to have **E** = 0 at any point *not* between the charges but on the line joining them, what are the necessary conditions and where is the point located? (*b*) Is it possible, for any arrangement of two point charges, to find *two* points (neither at infinity) at which **E** = 0; if so, under what conditions?

15. A thin nonconducting rod of finite length l carries a total charge q, spread uniformly along it. Show that E at point P on the perpendicular bisector in Fig. 27–20 is given by

$$E = \frac{q}{2\pi\epsilon_0 y}\frac{1}{\sqrt{l^2 + 4y^2}}.$$

Show that as $l \rightarrow \infty$ this result approaches that of Example 6.

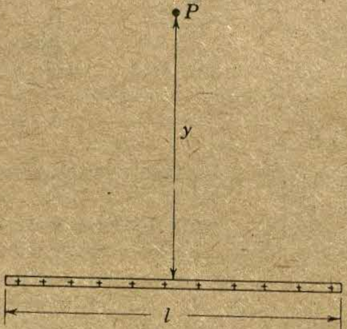


Fig. 27–20

16. A thin nonconducting rod is bent to form the arc of a circle of radius a and subtends an angle θ_0 at the center of the circle. A total charge q is spread uniformly along its length. Find the electric field strength at the center of the circle in terms of a, q, and θ_0.

17. A nonconducting hemispherical cup of inner radius a has a total charge q spread uniformly over its inner surface. Find the electric field at the center of curvature.

18. *Electric quadrupole.* Figure 27–21 shows a typical electric quadrupole. It consists of two dipoles whose effects at external points do not quite cancel. Show that the value of E on the axis of the quadrupole for points distant r from its center (assume $r \gg a$) is given by

$$E = \frac{3Q}{4\pi\epsilon_0 r^4}$$

where Q ($= 2qa^2$) is called the *quadrupole moment* of the charge distribution.

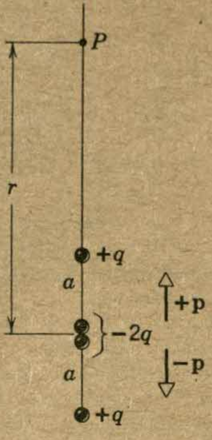


Fig. 27–21

19. An electron is constrained to move along the axis of the ring of charge in Example 5. Show that the electron can perform oscillations whose frequency is given by

$$\omega = \sqrt{\frac{eq}{4\pi\epsilon_0 m a^3}}.$$

This formula holds only for small oscillations, that is, for $x \ll a$ in Fig. 27–10. (Hint: Show that the motion is simple harmonic and use Eq. 15–11.)

20. For the ring of charge in Example 5, show that the maximum value of E occurs at $x = a/\sqrt{2}$.

21. Consider the ring of charge of Example 5. Suppose that the charge q is not distributed uniformly over the ring but that charge q_1 is distributed uniformly over half the circumference and charge q_2 is distributed uniformly over the other half. Let $q_1 + q_2 = q$. (*a*) Find the *component* of the electric field at any point on the axis directed *along* the axis and compare with the uniform case of Example 5. (*b*) Find the *component* of the electric field at any point on the axis *perpendicular* to the axis and compare with the uniform case of Example 5.

22. A thin circular disk of radius a is charged uniformly so as to have a charge per unit area of σ. Find the electric field on the axis of the disk at a distance r from the disk.

23. *Field due to an electric dipole.* Show that the components of **E** due to a dipole are given, at distant points, by

$$E_x = \frac{1}{4\pi\epsilon_0}\frac{3pxy}{(x^2 + y^2)^{5/2}}$$

$$E_y = \frac{1}{4\pi\epsilon_0}\frac{p(2y^2 - x^2)}{(x^2 + y^2)^{5/2}},$$

where x and y are coordinates of a point in Fig. 27–22. Show that this general result includes the special results of Eq. 27–8*b* and of Problem 10.

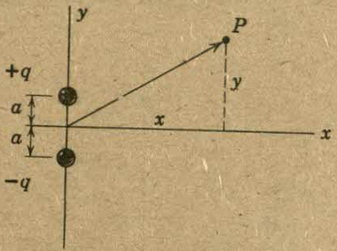


Fig. 27–22

24. What is the magnitude and direction of an electric field that will balance the weight of (a) an electron and (b) an alpha particle?

25. A particle having a charge of -2.0×10^{-9} coul is acted on by a downward electric force of 3.0×10^{-6} newton in a uniform electric field. (*a*) What is the strength of the electric field? (*b*) What is the magnitude and direction of the *electric* force exerted on a proton placed in this field? (*c*) What is the *gravitational* force on the proton? (*d*) What is the ratio of the electric to the gravitational forces in this case?

26. (*a*) What is the acceleration of an electron in a uniform electric field of 10^6 nt/coul? (*b*) How long would it take for the electron, starting from rest, to attain one-tenth the speed of light? (*c*) What considerations limit the applicability of Newtonian mechanics to such problems?

27. An electron moving with a speed of 5.0×10^8 cm/sec is shot parallel to an electric field of strength 1.0×10^3 nt/coul arranged so as to retard its motion. (*a*) How far will the electron travel in the field before coming (momentarily) to rest, and (*b*) how much time will elapse? (*c*) If the electric field ends abruptly after 0.8 cm, what fraction of its initial energy will the electron lose in traversing it?

28. An electron is projected as in Fig. 27–23 at a speed of 6.0×10^6 meters/sec and at an angle θ of 45°; $E = 2.0 \times 10^3$ nt/coul (directed upward), d = 2.0 cm, and l = 10.0 cm (*a*) Will the electron strike either of the plates? (*b*) If it strikes a plate, where does it do so?

29. *Dipole in a nonuniform field.* Derive an expression for dE/dz at a point midway between two equal positive charges, where z is the distance from one of the charges, measured along the line joining them. Would there be a force on a small dipole placed at this point, its axis being aligned with the z axis? Recall that $\mathbf{E} = 0$ at this point.

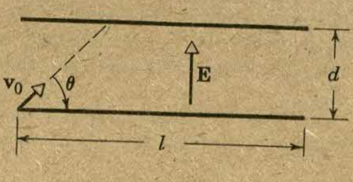


Fig. 27–23

30. *Oil drop experiment.* R. A. Millikan set up an apparatus (Fig. 27-24) in which a tiny, charged oil drop, placed in an electric field **E**, could be "balanced" by adjusting E until the electric force on the drop was equal and opposite to its weight. If the radius of the drop is 1.64×10^{-4} cm and E at balance is 1.92×10^{5} nt/coul, (*a*) what charge is on the drop? (*b*) Why did Millikan not try to balance electrons in his apparatus instead of oil drops? The density of the oil is 0.851 gm/cm^3. (Millikan first measured the electronic charge in this way. He measured the drop radius by observing the limiting speed that the drops attained when they fell in air with the electric field turned off. He charged the oil drops by irradiating them with bursts of X-rays.)

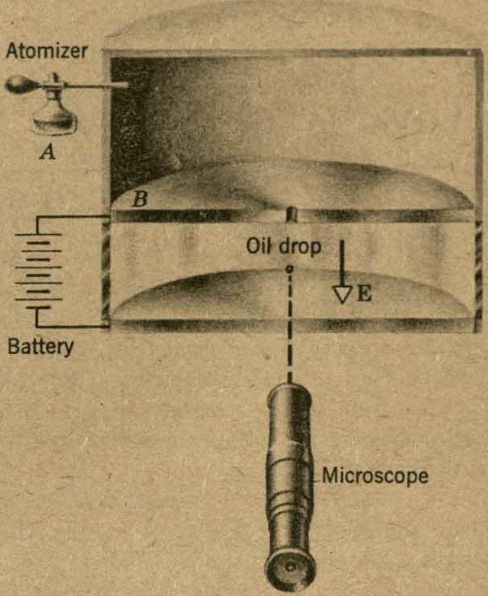


Fig. 27-24 Millikan's oil drop apparatus. Charged oil drops from atomizer A fall through the hole in plate B.

31. In a particular early run (1911), Millikan observed that the following measured charges, among others, appeared at different times on a single drop:

6.563×10^{-19} coul
8.204×10^{-19} coul
11.50×10^{-19} coul
13.13×10^{-19} coul
16.48×10^{-19} coul
18.08×10^{-19} coul
19.71×10^{-19} coul
22.89×10^{-19} coul
26.13×10^{-19} coul

What value for the elementary charge e can be deduced from these data?

32. An electric field **E** with an average magnitude of about 150 nt/coul points upward in the earth's atmosphere. We wish to "float" a sulfur sphere weighing 1.0 lb in this field by charging it. (*a*) What charge (sign and magnitude) must be used? (*b*) Why is the experiment not practical? Give a qualitative reason supported by a very rough numerical calculation to prove your point.

Gauss's Law

CHAPTER 28

28–1 Flux of the Electric Field

Flux (symbol Φ) is a property of any vector field; it refers to a hypothetical surface in the field, which may be closed or open. For a *flow field* the flux (Φ_v) is measured by the number of streamlines that cut through such a surface. For an *electric field* the flux (Φ_E) is measured by the number of lines of force that cut through such a surface.

For closed surfaces we shall see that Φ_E is positive if the lines of force point outward everywhere and negative if they point inward. Figure 28–1 shows two equal and opposite charges and their lines of force. Curves S_1, S_2, S_3, and S_4 are the intersections with the plane of the figure of four hypothetical closed surfaces. From the statement just given, Φ_E is positive for surface S_1 and negative for S_2. The flux of the electric field is important because Gauss's law, one of the four basic equations of electromagnetism (see Table 38–3), is expressed in terms of it. Although the concept of flux may seem a little abstract at first, the student will soon see its value in solving problems.

To define Φ_E precisely, consider Fig. 28–2, which shows an arbitrary closed surface immersed in an electric field. Let the surface be divided into elementary squares ΔS, each of which is small enough so that it may be considered to be plane. Such an element of area can be represented as a vector $\Delta \mathbf{S}$, whose magnitude is the area ΔS; the direction of $\Delta \mathbf{S}$ is taken as the *outward-drawn normal* to the surface.

At every square in Fig. 28–2 we can also construct an electric field vector $\mathbf{E}$. Since the squares have been taken to be arbitrarily small, $\mathbf{E}$ may be taken as constant for all points in a given square.

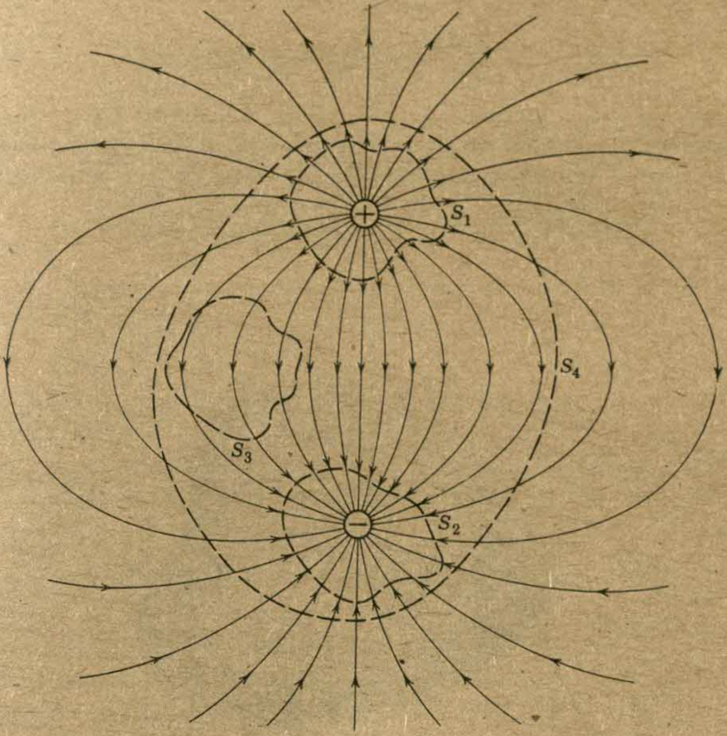


Fig. 28–1 Two equal and opposite charges. The dashed lines represent hypothetical closed surfaces.

The vectors **E** and $\Delta\mathbf{S}$ that characterize each square make an angle θ with each other. Figure 28–2*b* shows an enlarged view of the three squares on the surface of Fig. 28–2*a* marked x, y, and z. Note that at x, $\theta > 90°$; at y, $\theta = 90°$; and at z, $\theta < 90°$.

A semiquantitative definition of flux is

$$\Phi_E \cong \Sigma \mathbf{E}\cdot\Delta\mathbf{S}, \tag{28–1}$$

which instructs us to add up the scalar quantity $\mathbf{E}\cdot\Delta\mathbf{S}$ for all elements of area into which the surface has been divided. For points such as x in Fig. 28–2 the contribution to the flux is negative; at y it is zero and at z it is positive. Thus if **E** is everywhere outward, $\theta < 90°$, $\mathbf{E}\cdot\Delta\mathbf{S}$ will be positive, and Φ_E for the entire surface will be positive; see Fig. 28–1, surface S_1. If **E** is everywhere inward, $\theta > 90°$, $\mathbf{E}\cdot\Delta\mathbf{S}$ will be negative, and Φ_E for the surface will be negative; see Fig. 28–1, surface S_2. From Eq. 28-1 we see that the appropriate mks unit for Φ_E is the newton-meter^2/coul.

The exact definition of electric flux is found in the differential limit of Eq.

28–2 Gauss's Law

Gauss's law, which applies to any closed hypothetical surface (called a *Gaussian surface*), gives a connection between Φ_E for the surface and the net charge q enclosed by the surface. It is

$$\epsilon_0 \Phi_E = q \tag{28–3}$$

or, using Eq. 28–2,

$$\epsilon_0 \oint \mathbf{E} \cdot d\mathbf{S} = q. \tag{28–4}$$

The fact that Φ_E proves to be zero in Example 1 is predicted by Gauss's law because no charge is enclosed by the Gaussian surface in Fig. 28–3 ($q = 0$).

Note that q in Eq. 28–3 (or in Eq. 28–4) is the *net* charge, taking its algebraic sign into account. If a surface encloses equal and opposite charges, the flux Φ_E is zero. Charge outside the surface makes no contribution to the value of q, nor does the exact location of the inside charges affect this value.

Gauss's law can be used to evaluate **E** if the charge distribution is so symmetric that by proper choice of a Gaussian surface we can easily evaluate the integral in Eq. 28–4. Conversely, if **E** is known for all points on a given closed surface, Gauss's law can be used to compute the charge inside. If **E** has an outward component for every point on a closed surface, Φ_E, as Eq. 28–2 shows, will be positive and, from Eq. 28–4, there must be a net positive charge within the surface (see Fig. 28–1, surface S_1). If **E** has an inward component for every point on a closed surface, there must be a net negative charge within the surface (see Fig. 28–1, surface S_2). Surface S_3 in Fig. 28–1 encloses no charge, so that Gauss's law predicts that $\Phi_E = 0$. This is consistent with the fact that lines of **E** pass directly through surface S_3, the contribution to the integral on one side canceling that on the other. What would be the value of Φ_E for surface S_4 in Fig. 28–1, which encloses both charges?

28–3 Gauss's Law and Coulomb's Law

Coulomb's law can be deduced from Gauss's law and symmetry considerations. To do so, let us apply Gauss's law to an isolated point charge q as in Fig. 28–4. Although Gauss's law holds for any surface whatever, information can most readily be extracted for a spherical surface of radius r centered

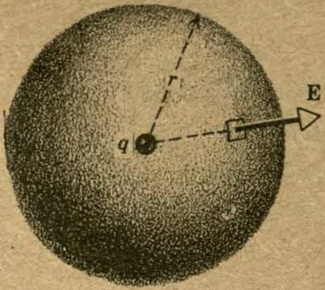


Fig. 28–4 A spherical Gaussian surface of radius r surrounding a point charge q.

on the charge. The advantage of this surface is that, from symmetry, **E** must be normal to it and must have the same (as yet unknown) magnitude for all points on the surface.

In Fig. 28-4 both **E** and $d\mathbf{S}$ at any point on the Gaussian surface are directed radially outward. The angle between them is zero and the quantity $\mathbf{E}\cdot d\mathbf{S}$ becomes simply $E\,dS$. Gauss's law (Eq. 28-4) thus reduces to

$$\epsilon_0 \oint \mathbf{E}\cdot d\mathbf{S} = \epsilon_0 \oint E\,dS = q.$$

Because E is constant for all points on the sphere, it can be factored from inside the integral sign, leaving

$$\epsilon_0 E \oint dS = q,$$

where the integral is simply the area of the sphere.* This equation gives

$$\epsilon_0 E(4\pi r^2) = q$$

or

$$E = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2}. \tag{28-5}$$

Equation 28-5 gives the magnitude of the electric field strength **E** at any point a distance r from an isolated point charge q. The direction of **E** is already known from symmetry.

Let us put a second point charge q_0 at the point at which **E** is calculated. The magnitude of the force that acts on it (see Eq. 27-2) is

$$F = Eq_0.$$

Combining with Eq. 28-5 gives

$$F = \frac{1}{4\pi\epsilon_0}\frac{qq_0}{r^2},$$

which is precisely Coulomb's law. Thus we have deduced Coulomb's law from Gauss's law and considerations of symmetry.

Gauss's law is one of the fundamental equations of electromagnetic theory and is displayed in Table 38-3 as one of Maxwell's equations. Coulomb's law is not listed in that table because, as we have just proved, it can be deduced from Gauss's law and from simple assumptions about the symmetry of **E** due to a point charge.

It is interesting to note that writing the proportionality constant in Coulomb's law as $1/4\pi\epsilon_0$ (see Eq. 26-3) permits a particularly simple form for Gauss's law (Eq. 28-3). If we had written the Coulomb law constant simply as k, Gauss's law would have to be written as $(1/4\pi k)\Phi_E = q$. We prefer to leave the factor 4π in Coulomb's law so that it will not appear in Gauss's law or in other much used relations that will be derived later.

---

* The usefulness of Gauss's law depends on our ability to find a surface over which, from symmetry, both E and θ (see Fig. 28-2) have constant values. Then $E\cos\theta$ can be factored out of the integral and E can be found simply, as in this example.

28-4 An Insulated Conductor

Gauss's law can be used to make an important prediction, namely: *An excess charge, placed on an insulated conductor, resides entirely on its outer surface.* This hypothesis was shown to be true by experiment (see Section 28-5) before either Gauss's law or Coulomb's law were advanced. Indeed, the experimental proof of the hypothesis is the experimental foundation upon which both laws rest: We have already pointed out that Coulomb's torsion balance experiments, although direct and convincing, are not capable of great accuracy. In showing that the italicized hypothesis is predicted by Gauss's law, we are simply reversing the historical situation.

Figure 28-5 is a cross section of an insulated conductor of arbitrary shape carrying an excess charge q. The dashed lines show a Gaussian surface that lies a small distance below the actual surface of the conductor. Although the Gaussian surface can be as close to the actual surface as we wish, it is important to keep in mind that the Gaussian surface is *inside* the conductor.

When an excess charge is placed at random on an insulated conductor, it will set up electric fields inside the conductor. These fields act on the charge carriers of the conductor and cause them to move, that is, they set up internal currents. These currents redistribute the excess charge in such a way that the internal electric fields are automatically reduced in magnitude. Eventually the electric fields inside the conductor become zero everywhere, the currents automatically stop, and electrostatic conditions prevail. This redistribution of charge normally takes place in a time that is negligible for most purposes. What can be said about the distribution of the excess charge when such electrostatic conditions have been achieved?

If, at electrostatic equilibrium, **E** is zero everywhere inside the conductor, it must be zero for every point on the Gaussian surface. This means that the flux Φ_E for this surface must be zero. Gauss's law then predicts (see Eq. 28-3) that there must be no net charge inside the Gaussian surface. If the excess charge q is not *inside* this surface, it can only be *outside* it, that is, *it must be on the actual surface of the conductor.*

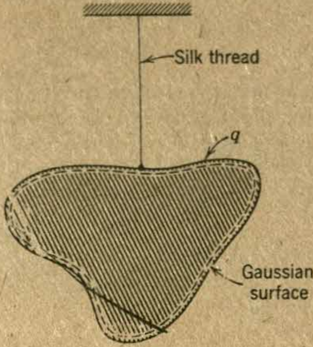


Fig. 28-5 An insulated conductor.

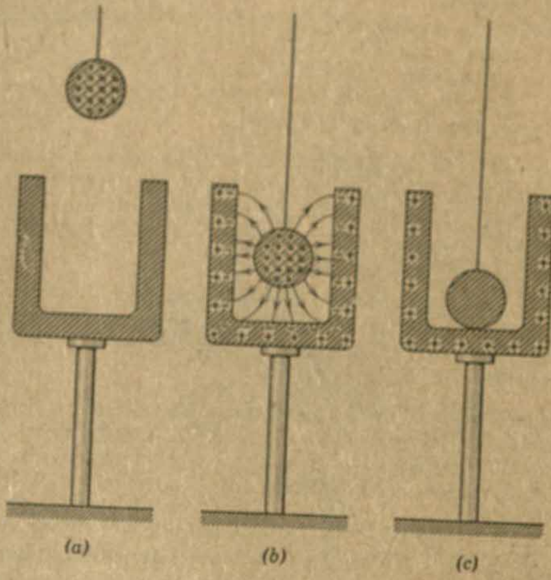


Fig. 28–6 The *entire* charge on the ball is transferred to the *outside* of the can. This statement and the discussion of the first paragraph of Section 28–5 are strictly correct only if the can is provided with a conducting lid which can be closed after the ball is inserted.

28–5 Experimental Proof of Gauss's and Coulomb's Laws

Let us turn to the experiments that prove that the hypothesis of Section 28–4. is true. For a simple test, charge a metal ball and lower it with a silk thread deep into a metal can as in Fig. 28–6. Touch the ball to the inside of the can; when the ball is removed from the can, all its charge will have vanished. When the metal ball touches the can, the ball and can together form an "insulated conductor" to which the hypothesis of Section 28–4 applies. That the charge moves entirely to the outside surface of the can can be shown by touching a small insulated metal object to the can; only on the *outside* of the can will it be possible to pick up a charge.

Benjamin Franklin seems to have been the first to notice that there can be no charge inside an insulated metal can. In 1755 he wrote to a friend:

> I electrified a silver pint cann, on an electric stand, and then lowered into it a cork-ball, of about an inch diameter, hanging by a silk string, till the cork touched the bottom of the cann. The cork was not attracted to the inside of the cann as it would have been to the outside, and though it touched the bottom, yet when drawn out, it was not found to be electrified by that touch, as it would have been by touching the outside. The fact is singular. You require the reason; I do not know it. . . .

About ten years later Franklin recommended this "singular fact" to the attention of his friend Joseph Priestley (1733–1804). In 1767 (about twenty years before Coulomb's experiments) Priestley checked Franklin's observation and, with remarkable insight, realized that the inverse square law of force followed from it. This is

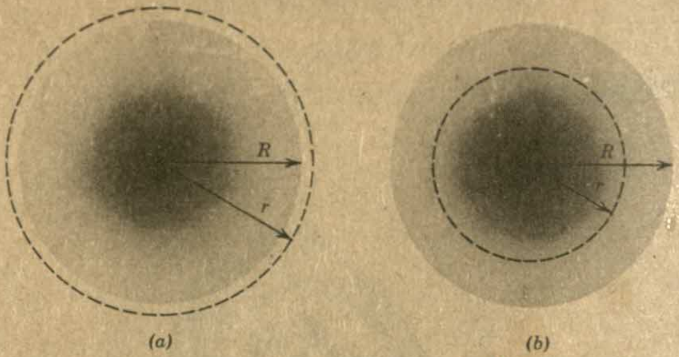


Fig. 28–8 Example 2. A spherically symmetric charge distribution, showing two Gaussian surfaces. The density of charge, as the shading suggests, varies with distance from the center but not with direction.

tribution of charge, the electric field has the value that it would have if the charge were concentrated at its center. This reminds us that a sphere of mass m behaves gravitationally, for outside points, as if the mass were concentrated at its center. At the root of this similarity lies the fact that both Coulomb's law and the law of gravitation are inverse square laws. The gravitational case was proved in detail in Section 16–6; the proof using Gauss's law in the electrostatic case is certainly much simpler.

Figure 28–8*b* shows a spherical Gaussian surface of radius r drawn *inside* the charge distribution. Gauss's law (Eq. 28–4) gives

$$\epsilon_0 \oint \mathbf{E} \cdot d\mathbf{S} = \epsilon_0 E(4\pi r^2) = q'$$

or

$$E = \frac{1}{4\pi\epsilon_0} \frac{q'}{r^2},$$

in which q' is that part of q contained within the sphere of radius r. The part of q that lies outside this sphere makes no contribution to **E** at radius r. This corresponds, in the gravitational case (Section 16–6), to the fact that a spherical shell of matter exerts no gravitational force on a body inside it.

An interesting special case of a spherically symmetric charge distribution is a uniform sphere of charge. For such a sphere, which would be suggested by uniform shading in Fig. 28–8, the charge density ρ would have a constant value for all points within a sphere of radius R and would be zero for all points outside this sphere. For points inside such a uniform sphere of charge we can put

$$q' = q\,\frac{\frac{4}{3}\pi r^3}{\frac{4}{3}\pi R^3}$$

or

$$q' = q\left(\frac{r}{R}\right)^3$$

where $\frac{4}{3}\pi R^3$ is the volume of the spherical charge distribution. The expression for E then becomes

$$E = \frac{1}{4\pi\epsilon_0} \frac{qr}{R^3}. \tag{28–6}$$

This equation becomes zero, as it should, for $r = 0$. Note that Eqs. 28-5 and 28-6 give the same result, as they must, for points on the surface of the charge distribution (that is, if $r = R$). Note that Eq. 28-6 does not apply to the charge distribution of Fig. 28-8*b* because the charge density, suggested by the shading, is *not* constant in that case.

Example 3. *The Thomson atom model.* At one time the positive charge in the atom was thought to be distributed uniformly throughout a sphere with a radius of about 1.0×10^{-10} meter, that is, throughout the entire atom. Calculate the electric field strength at the surface of a gold atom ($Z = 79$) on this (erroneous) assumption. Neglect the effect of the electrons.

The positive charge of the atom is Ze or $(79)(1.6 \times 10^{-19}\text{ coul})$. Equation 28-5 yields, for E at the surface,

$$\begin{aligned} E &= \frac{1}{4\pi\epsilon_0}\frac{q}{r^2} \\ &= \frac{(9.0 \times 10^{9}\text{ nt-m}^2/\text{coul}^2)(79)(1.6 \times 10^{-19}\text{ coul})}{(1.0 \times 10^{-10}\text{ meter})^2} \\ &= 1.1 \times 10^{13}\text{ nt/coul.} \end{aligned}$$

Figure 28-9 is a plot of E as a function of distance from the center of the atom, using Eqs. 28-5 and 28-6. We see that E has its maximum value on the surface and decreases linearly to zero at the center (see Eq. 28-6). Outside the sphere E decreases as the inverse square of the distance (see Eq. 28-5).

Example 4. *The Rutherford, or nuclear, atom.* We shall see in Section 28-7 that the positive charge of the atom is *not* spread uniformly throughout the atom (see Example 3) but is concentrated in a small region (the *nucleus*) at the center of the atom. For gold the radius of the nucleus is about 6.9×10^{-15} meter. What is the electric field strength at the nuclear surface? Again neglect effects associated with the atomic electrons.

The problem is the same as that of Example 3, except that the radius is much smaller. This will make the electric field strength at the surface larger, in proportion to the ratio of the squares of the radii. Thus

$$\begin{aligned} E &= (1.1 \times 10^{13}\text{ nt/coul})\frac{(1.0 \times 10^{-10}\text{ meter})^2}{(6.9 \times 10^{-15}\text{ meter})^2} \\ &= 2.3 \times 10^{21}\text{ nt/coul.} \end{aligned}$$

This is an enormous electric field, much stronger than could be produced and maintained in the laboratory. It is about 10^8 times as large as the field calculated in Example 3.

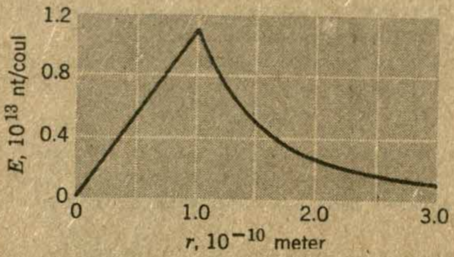


Fig. 28-9 Example 3. The electric field due to the positive charge in a gold atom, according to the (erroneous) Thomson model.

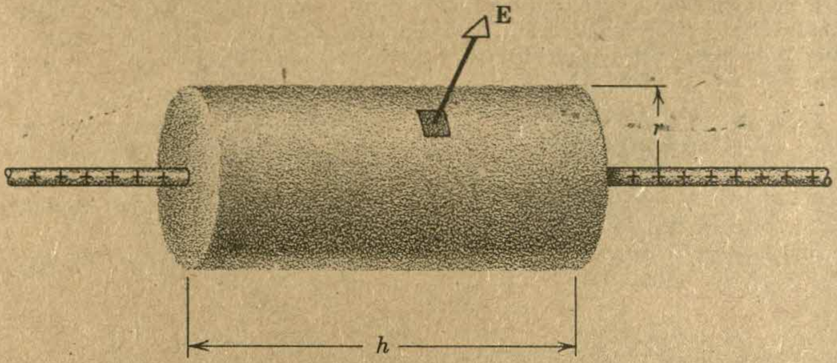


Fig. 28–10 Example 5. An infinite rod of charge, showing a cylindrical Gaussian surface.

Example 5. *Line of charge.* Figure 28–10 shows a section of an infinite rod of charge, the *linear charge density* λ (that is, the charge per unit length, measured in coul/meter) being constant for all points on the line. Find an expression for E at a distance r from the line.

From symmetry, **E** due to a uniform linear charge can only be radially directed. As a Gaussian surface we choose a circular cylinder of radius r and length h, closed at each end by plane caps normal to the axis. E is constant over the cylindrical surface and the flux of **E** through this surface is $E(2\pi rh)$ where $2\pi rh$ is the area of the surface. There is no flux through the circular caps because **E** here lies in the surface at every point.

The charge enclosed by the Gaussian surface of Fig. 28–10 is λh. Gauss's law (Eq. 28–4),

$$\epsilon_0 \oint \mathbf{E} \cdot d\mathbf{S} = q,$$

then becomes

$$\epsilon_0 E(2\pi rh) = \lambda h,$$

whence

$$E = \frac{\lambda}{2\pi\epsilon_0 r}. \tag{28-7}$$

The direction of **E** is radially outward for a line of positive charge.

Note how much simpler the solution using Gauss's law is than that using integration methods, as in Example 6, Chapter 27. Note too that the solution using Gauss's law is possible only if we choose our Gaussian surface to take full advantage of the radial symmetry of the electric field set up by a long line of charge. We are free to choose any surface, such as a cube or a sphere, for a Gaussian surface. Even though Gauss's law holds for all such surfaces, they are not all useful for the problem at hand; only the cylindrical surface of Fig. 28–10 is appropriate in this case.

Gauss's law has the property that it provides a useful technique for calculation only in problems that have a certain degree of symmetry, but in these problems the solutions are strikingly simple.

Example 6. *A sheet of charge.* Figure 28–11 shows a portion of a thin, *nonconducting*, infinite sheet of charge, the *surface charge density* σ (that is, the charge per unit area, measured in coul/meter^2) being constant. What is **E** at a distance r in front of the plane?

A convenient Gaussian surface is a "pill box" of cross-sectional area A and height $2r$, arranged to pierce the plane as shown From symmetry, **E** points at right angles

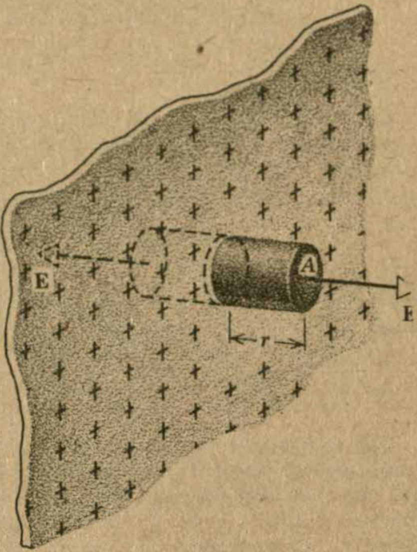


Fig. 28–11 Example 6. An infinite sheet of charge pierced by a cylindrical Gaussian surface. The cross section of the cylinder need not be circular, as shown, but can have an arbitrary shape.

to the end caps and away from the plane. Since **E** does not pierce the cylindrical surface, there is no contribution to the flux from this source. Thus Gauss's law,

$$\epsilon_0 \oint \mathbf{E} \cdot d\mathbf{S} = q$$

becomes

$$\epsilon_0(EA + EA) = \sigma A$$

where σA is the enclosed charge. This gives

$$E = \frac{\sigma}{2\epsilon_0}. \tag{28-8}$$

Note that E is the same for all points on each side of the plane; compare Fig. 27–2. Although an infinite sheet of charge cannot exist physically, this derivation is still useful in that Eq. 28–8 yields substantially correct results for real (not infinite) charge sheets if we consider only points not near the edges whose distance from the sheet is small compared to the dimensions of the sheet.

Example 7. *A charged conductor.* Figure 28–12 shows a *conductor* carrying on its surface a charge whose surface charge density at any point is σ; in general σ will vary from point to point. What is **E** for points a short distance above the surface?

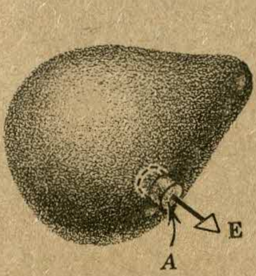


Fig. 28–12 Example 7. A charged insulated conductor showing a Gaussian surface. The cross section of the surface need not be circular, as shown, but can have an arbitrary shape.

The direction of **E** *for points close to the surface* is at right angles to the surface, pointing away from the surface if the charge is positive. If **E** were *not* normal to the surface, it would have a component lying in the surface. Such a component would act on the charge carriers in the conductor and set up surface currents. Since there are no such currents under the assumed electrostatic conditions, **E** must be normal to the surface.

The magnitude of **E** can be found from Gauss's law using a small flat "pill box" of cross section A as a Gaussian surface. Since **E** equals zero everywhere inside the conductor (see Section 28–4), the only contribution to Φ_E is through the plane cap of area A that lies outside the conductor. Gauss's law

$$\epsilon_0\oint \mathbf{E}\cdot d\mathbf{S} = q$$

becomes

$$\epsilon_0(EA) = \sigma A$$

where σA is the net charge within the Gaussian surface. This yields

$$E = \frac{\sigma}{\epsilon_0}. \qquad (28\text{–}9)$$

Comparison with Eq. 28–8 shows that the electric field is *twice as great* near a conductor carrying a charge whose surface charge density is σ as that near a nonconducting sheet with the same surface charge density. The student should compare the Gaussian surfaces in Figs. 28–11 and 28–12 carefully. In Fig. 28–11 lines of force leave the surface through *each* end cap, an electric field existing on *both* sides of the sheet. In Fig. 28–12 the lines of force leave only through the *outside* end cap, the inner end cap being inside the conductor where no electric field exists. If we assume the same surface charge density and cross-sectional area A for the two Gaussian surfaces, the enclosed charge ($= \sigma A$) will be the same. Since, from Gauss's law, the flux Φ_E must then be the same in each case, it follows that E ($= \Phi_E/A$) must be twice as large in Fig. 28–12 as in Fig. 28–11. It is helpful to note that in Fig. 28–11 half the flux emerges from one side of the surface and half from the other, whereas in Fig. 28–12 all the flux emerges from the outside surface. ◄

28–7 The Nuclear Model of the Atom

Ernest Rutherford (1871–1937) was first led, in 1911, to assume that the atomic nucleus existed when he tried to interpret some experiments carried out at the University of Manchester by his collaborators H. Geiger and E. Marsden.* The results of Examples 3 and 4 played an important part in Rutherford's analysis of these experiments.

These workers, at Rutherford's suggestion, allowed a beam of α-particles † to strike and be deflected by a thin film of a heavy element such as gold. They counted the number of particles deflected through various angles ϕ. Figure 28–13 shows the experimental setup schematically. Figure 28–14 shows the paths taken by typical α-particles as they scatter from a gold atom; the angles ϕ through which the α-particles are deflected range from 0 to 180° as the character of the collision varies from "grazing" to "head-on."

---

* See "The Birth of the Nuclear Atom," E. N. da C. Andrade, *Scientific American*, November 1956. See also Example 5, Chapter 10.

† α-Particles are helium nuclei that are emitted spontaneously by some radioactive materials such as radium. They move with speeds of the order of one-thirtieth that of light when so emitted.

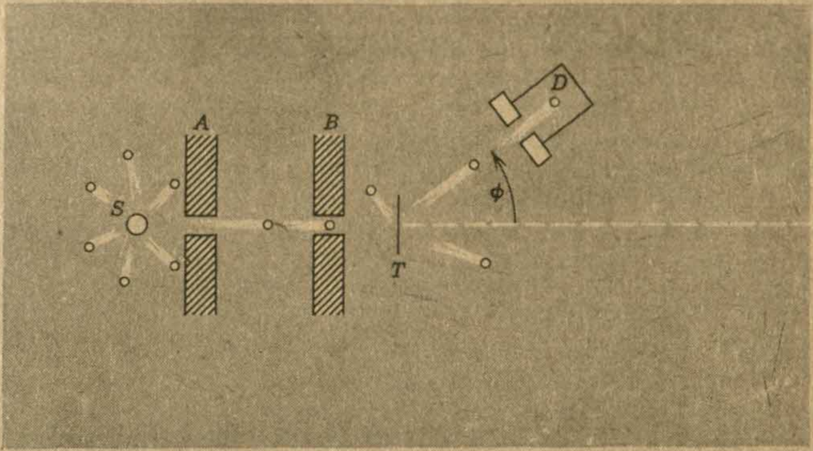


Fig. 28-13 Experimental arrangement for studying the scattering of α-particles. Particles from radioactive source S are allowed to fall on a thin metal "target" T; α-particles scattered by the target through an (adjustable) angle ϕ are counted by detector D.

The electrons in the gold atom, being so light, have almost no effect on the motion of an oncoming α-particle; the electrons are themselves strongly deflected, just as a swarm of insects would be by a stone hurled through them. Any deflection of the α-particle must be caused by the repulsive action of the positive charge of the gold atom, which is known to possess most of the mass of the atom.

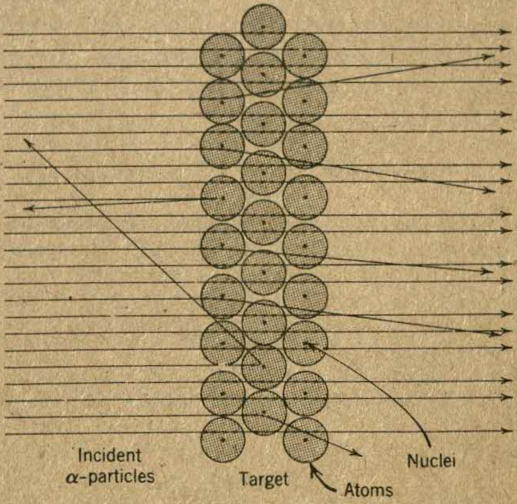


Fig. 28-14 The deflection of the incident α-particles depends on the nature of the nuclear collision. (From Andrade, *Scientific American*, November 1956.)

At the time of these experiments most physicists believed in the so-called "plum pudding" model of the atom that had been suggested by J. J. Thomson (1856–1940). In this view (see Example 3) the positive charge of the atom was thought to be spread out through the whole atom, that is, through a spherical volume of radius about 10^{-10} meter. The electrons were thought to vibrate about fixed centers inside this sphere.

Rutherford showed that this model of the atom was not consistent with the α-scattering experiments and proposed instead the nuclear model of the atom that we now accept. Here the positive charge is confined to a very much smaller sphere whose radius is about 10^{-14} meter (the *nucleus*). The electrons move around this nucleus and occupy a roughly spherical volume of radius about 10^{-10} meter. This brilliant deduction by Rutherford laid the foundation for modern atomic and nuclear physics.

The feature of the α-scattering experiments that attracted Rutherford's attention at once was that a few α-particles are deflected through very large angles, up to 180°. To scientists accustomed to thinking in terms of the "plum pudding" model, this was a very surprising result. In Rutherford's words: "It was quite the most incredible event that ever happened to me in my life. It was almost as incredible as if you had fired a 15-inch shell at a piece of tissue paper and it came back and hit you."

The α-particle must pass through a region in which the electric field strength is very high indeed in order to be deflected so strongly.* Example 3 shows that, in Thomson's model, the maximum electric field strength is 1.1×10^{13} nt/coul. Compare this with the value calculated in Example 4 for a point on the surface of a gold nucleus (2.3×10^{21} nt/coul). Thus the deflecting force acting on an α-particle can be up to 10^8 times as great if the positive charge of the atom is compressed into a small enough region (the nucleus) at the center of the atoms. Rutherford made his hypothesis about the existence of nuclei only after a much more detailed mathematical analysis than that given here.

QUESTIONS

1. A point charge is placed at the center of a spherical Gaussian surface. Is Φ_E changed (*a*) if the surface is replaced by a cube of the same volume, (*b*) if the sphere is replaced by a cube of one-tenth the volume, (*c*) if the charge is moved off-center in the original sphere, still remaining inside, (*d*) if the charge is moved just outside the original sphere, (*e*) if a second charge is placed near, and outside, the original sphere, and (*f*) if a second charge is placed inside the Gaussian surface?

2. By analogy with Φ_E, how would you define the flux Φ_g of a gravitational field? What is the flux of the earth's gravitational field through the boundaries of a room, assumed to contain no matter?

3. In Gauss's law,

$$\epsilon_0 \oint \mathbf{E} \cdot d\mathbf{S} = q,$$

is **E** the electric field intensity attributable to the charge q?

---

* The chance that a big deflection can result from the combined effects of many small deflections can be shown to be very small.

4. Show that Eq. 18–3 illustrates what might be called *Gauss's law for incompressible fluids*, or

$$\Phi_v = \oint \mathbf{v} \cdot d\mathbf{S} = 0.$$

5. A surface encloses an electric dipole. What can you say about Φ_E for this surface?

6. Suppose that a Gaussian surface encloses no net charge. Does Gauss's law require that **E** equal zero for all points on the surface? Is the converse of this statement true, that is, if **E** equals zero everywhere on the surface, does Gauss's law require that there be no net charge inside?

7. Would Gauss's law hold if the exponent in Coulomb's law were not exactly two?

8. Does Gauss's law, as applied in Section 28–4, require that all the conduction electrons in an insulated conductor reside on the surface?

9. In Section 28–4 we assumed that **E** equals zero everywhere inside a conductor. However, there are certainly very large electric fields inside the conductor, at points close to the electrons or to the nuclei. Does this invalidate the proof of Section 28–4?

10. It is sometimes said that excess charge resides entirely on the outer surface of a conductor because like charges repel and try to get as far away as possible from one another. Comment on this plausibility argument.

11. Is Gauss's law useful in calculating the field due to three equal charges located at the corners of an equilateral triangle? Explain.

12. The use of line, surface, and volume densities of charge to calculate the charge contained in an element of a charged object implies a continuous distribution of charge, whereas, in fact, charge on the microscopic scale is discontinuous. How, then, is this procedure justified?

13. Is **E** necessarily zero inside a charged rubber balloon if the balloon is (*a*) spherical or (*b*) sausage-shaped? For each shape assume the charge to be distributed uniformly over the surface.

14. A spherical rubber balloon carries a charge that is uniformly distributed over its surface. How does E vary for points (*a*) inside the balloon, (*b*) at the surface of the balloon, and (*c*) outside the balloon, as the balloon is blown up?

15. As you penetrate a uniform sphere of charge, E should decrease because less charge lies inside a sphere drawn through the observation point. On the other hand, E should increase because you are closer to the center of this charge. Which effect predominates and why?

16. Given a spherically symmetric charge distribution (not of uniform density of charge), is E necessarily a maximum at the surface? Comment on various possibilities.

17. An atom is normally *electrically neutral*. Why then should an α-particle be deflected by the atom under any circumstances?

18. If an α-particle, fired at a gold nucleus, is deflected through 135°, can you conclude (*a*) that any force has acted on the α-particle or (*b*) that any net work has been done on it?

PROBLEMS

1. Calculate Φ_E through a hemisphere of radius R. The field of **E** is uniform and is parallel to the axis of the hemisphere.

2. In Example 1 compute Φ_E for the cylinder if it is turned so that its axis is perpendicular to the electric field. Do not use Gauss's law.

3. A plane surface of area A is inclined so that its normal makes an angle θ with a uniform field of **E**. Calculate Φ_E for this surface.

4. A point charge of 1.0×10^{-6} coul is at the center of a cubical Gaussian surface 0.50 meter on edge. What is Φ_E for the surface?

5. Charge on an originally uncharged insulated conductor is separated by holding a positively charged rod nearby, as in Fig. 28–15. What can you learn from Gauss's law about the flux for the five Gaussian surfaces shown? The induced negative charge on the conductor is equal to the positive charge on the rod.

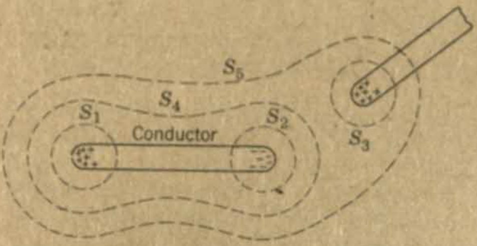


Fig. 28–15

6. "Gauss's law for gravitation" is

$$\frac{1}{4\pi G}\Phi_g = \frac{1}{4\pi G}\oint \mathbf{g}\cdot d\mathbf{S} = m,$$

where m is the enclosed mass and G is the universal gravitation constant (Section 16–3). Derive Newton's law of gravitation from this.

7. Figure 28–16 shows a point charge of 1.0×10^{-7} coul at the center of a spherical cavity of radius 3.0 cm in a piece of metal. Use Gauss's law to find the electric field at point a, halfway from the center to the surface, and at point b.

8. An uncharged spherical thin metallic shell has a point charge q at its center. Give expressions for the electric field (*a*) inside the shell, and (*b*) outside the shell, using Gauss's law. (*c*) Has the shell any effect on the field due to q? (*d*) Has the presence of q any effect on the shell? (*e*) If a second point charge is held outside the shell, does this outside charge experience a force? (*f*) Does the inside charge experience a force? (*g*) Is there a contradiction with Newton's third law here?

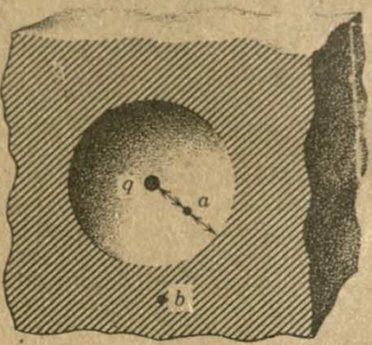


Fig. 28–16

9. Two large nonconducting sheets of positive charge face each other as in Fig. 28–17. What is **E** at points (*a*) to the left of the sheets, (*b*) between them, and (*c*) to the right of the sheets? Assume the same surface charge density σ for each sheet. Consider only points not near the edges whose distance from the sheets is small compared to the dimensions of the sheet. (Hint: **E** at any point is the vector sum of the separate electric field strengths set up by each sheet.)

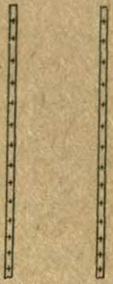


Fig. 28–17

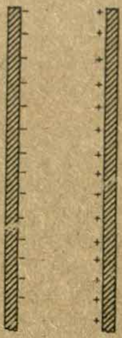


Fig. 28–18

10. Two large metal plates face each other as in Fig. 28-18 and carry charges with surface charge density $+\sigma$ and $-\sigma$, respectively, on their inner surfaces. What is **E** at points (*a*) to the left of the sheets, (*b*) between them, and (*c*) to the right of the sheets. Consider only points not near the edges whose distance from the sheets is small compared to the dimensions of the sheet.

11. Two large metal plates of area 1.0 meter^2 face each other. They are 5.0 cm apart and carry equal and opposite charges on their inner surfaces. If E between the plates is 55 nt/coul, what is the charge on the plates? Neglect edge effects. See Problem 10.

12. A thin-walled metal sphere has a radius of 25 cm and carries a charge of 2.0×10^{-7} coul. Find E for a point (*a*) inside the sphere, (*b*) just outside the sphere, and (*c*) 3.0 meters from the center of the sphere.

13. A 100-ev electron is fired directly toward a large metal plate that has a surface charge density of -2.0×10^{-6} coul/meter^2. From what distance must the electron be fired if it is to just fail to strike the plate?

14. Charge is distributed uniformly throughout an infinitely long cylinder of radius R. Show that E at a distance r from the cylinder axis ($r < R$) is given by

$$E = \frac{\rho r}{2\epsilon_0}$$

where ρ is the density of charge (coul/meter^3). What result do you expect for $r > R$?

15. Figure 28-19 shows a spherical nonconducting shell of charge of uniform density ρ coul/meter^3). Plot E for distances r from the center of the shell ranging from zero to 30 cm. Assume that $\rho = 1.0 \times 10^{-6}$ coul/meter^3, $a = 10$ cm, and $b = 20$ cm.

16. Figure 28-20 shows a section through a long, thin-walled metal tube of radius R, carrying a charge per unit length λ on its surface. Derive expressions for E for various distances r from the tube axis, considering both $r > R$ and $r < R$. Plot your results for the range $r = 0$ to $r = 5$ cm, assuming that $\lambda = 2.0 \times 10^{-8}$ coul/meter and $R = 3.0$ cm.

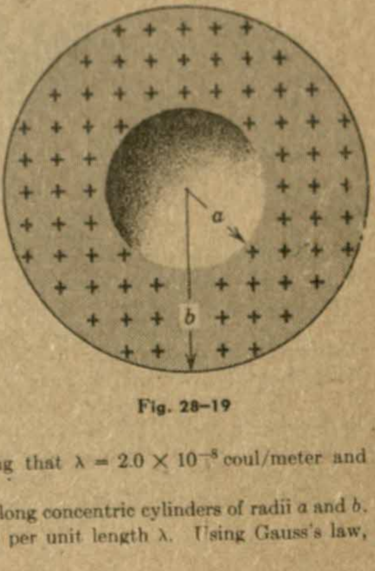


Fig. 28-19

17. Figure 28-21 shows a section through two long concentric cylinders of radii a and b. The cylinders carry equal and opposite charges per unit length λ. Using Gauss's law,

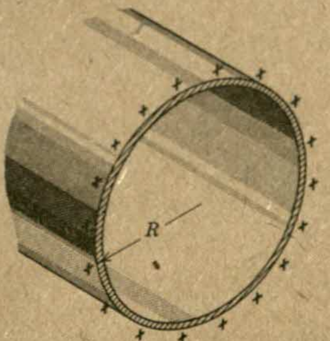


Fig. 28-20

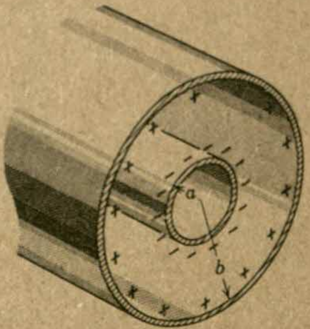


Fig. 28-21

prove (*a*) that $E = 0$ for $r > b$ and for $r < a$ and (*b*) that between the cylinders E is given by

$$E = \frac{1}{2\pi\epsilon_0}\frac{\lambda}{r}.$$

18. In Problem 17 a positron revolves in a circular path of radius r, between and concentric with the cylinders. What must be its kinetic energy K? Assume $a = 2.0$ cm, $b = 3.0$ cm, and $\lambda = 3.0 \times 10^{-8}$ coul/meter.

19. A long conducting cylinder carrying a total charge $+q$ is surrounded by a conducting cylindrical shell of total charge $-2q$ as shown in cross section in Fig. 28–22. Use Gauss's law to find (*a*) the electric field strength at points outside the conducting shell, (*b*) the distribution of the charge on the conducting shell, and (*c*) the electric field strength in the region between the cylinders. List the assumptions made in arriving at your answers.

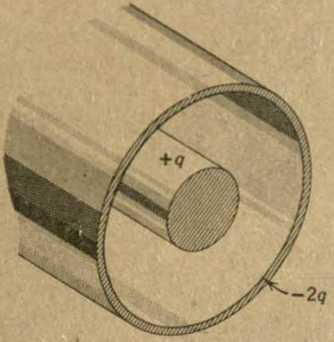


Fig. 28–22

20. A thin metallic spherical shell of radius a carries a charge q_a. Concentric with it is another thin metallic spherical shell of radius b ($b > a$) carrying a charge q_b. Use Gauss's law to find the electric field strength at radial points r where (*a*) $r < a$; (*b*) $a < r < b$; (*c*) $r > b$. (*d*) Discuss the criterion one would use to determine how the charges are distributed on the inner and outer surfaces of each shell.

21. A small sphere whose mass m is 1.0×10^{-3} gm carries a charge q of 2.0×10^{-8} coul. It hangs from a silk thread which makes an angle of 30° with a large, charged conducting sheet as in Fig. 28–23. Calculate the surface charge density σ for the sheet.

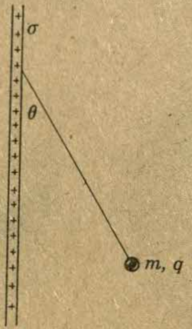


Fig. 28–23

22. Equation 28-9 ($E = \sigma/\epsilon_0$) gives the electric field at points near a charged conducting surface. Show that this equation leads to a familiar result when applied to a conducting sphere of radius r, carrying a charge q.

23. An α-particle, approaching the surface of a nucleus of gold, is a distance equal to one nuclear radius (6.9×10^{-15} meter) away from that surface. What are the forces on the α-particle and its acceleration at that point? The mass of the α-particle, which may be treated here as a point, is 6.7×10^{-27} kg.

24. A gold foil used in a Rutherford scattering experiment is 3×10^{-5} cm thick. (*a*) What fraction of its surface area is "blocked out" by gold nuclei, assuming a nuclear radius of 6.9×10^{-15} meter? Assume that no nucleus is screened by any other. (*b*) What fraction of the volume of the foil is occupied by the nuclei? (*c*) What fills all the rest of the space in the foil?

25. The electric field components in Fig. 28-24 are $E_x = bx^{1/2}$, $E_y = E_z = 0$, in which $b = 800$ nt/coul-m$^{1/2}$. Calculate (*a*) the flux Φ_E through the cube and (*b*) the charge within the cube. Assume that $a = 10$ cm.

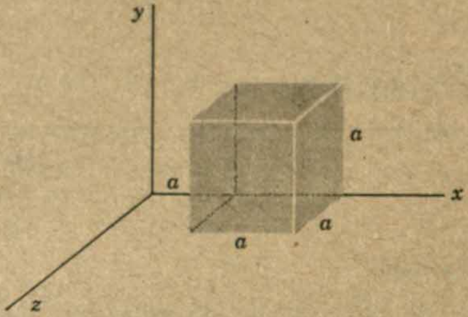


Fig. 28-24

Electric Potential

CHAPTER 29

29–1 Electric Potential

The electric field around a charged rod can be described not only by a (vector) electric field strength **E** but also by a scalar quantity, the *electric potential* V. These quantities are intimately related, and often it is only a matter of convenience which is used in a given problem.

To find the *electric potential difference* between two points A and B in an electric field, we move a test charge q_0 from A to B, always keeping it in equilibrium, and we measure the work W_{AB} that must be done by the agent moving the charge. The electric potential difference * is defined from

$$V_B - V_A = \frac{W_{AB}}{q_0}. \qquad (29\text{–}1)$$

The work W_{AB} may be (*a*) positive, (*b*) negative, or (*c*) zero. In these cases the electric potential at B will be (*a*) higher, (*b*) lower, or (*c*) the same as the electric potential at A.

The mks unit of potential difference that follows from Eq. 29–1 is the joule/coul. This combination occurs so often that a special unit, the *volt*, is used to represent it; that is,

$$1 \text{ volt} = 1 \text{ joule/coul.}$$

---

* This definition of potential difference, though conceptually sound and suitable for our present purpose, is rarely carried out in practice because of technical difficulties. Equivalent and more technically feasible methods are usually adopted.

Usually point A is chosen to be at a large (strictly an infinite) distance from all charges, and the electric potential V_A at this infinite distance is arbitrarily taken as zero. This allows us to define the *electric potential at a point.* Putting $V_A = 0$ in Eq. 29-1 and dropping the subscripts leads to

$$V = \frac{W}{q_0}, \tag{29-2}$$

where W is the work that an external agent must do to move the test charge q_0 from infinity to the point in question. The student should keep in mind that *potential differences* are of fundamental concern and that Eq. 29-2 depends on the arbitrary assignment of the value zero to the potential V_A at the reference position (infinity); this reference potential could equally well have been chosen as any other value, say -100 volts. Similarly any other agreed-upon point could be chosen as a reference position. In many circuit problems the *earth* is taken as a reference of potential and assigned the value zero.

Bearing in mind the assumptions made about the reference position, we see from Eq. 29-2 that V near an isolated positive charge is positive because positive work must be done by an outside agent to push a (positive) test charge in from infinity. Similarly, the potential near an isolated negative charge is negative because an outside agent must exert a restraining force on (that is, must do negative work on) a (positive) test charge as it comes in from infinity. Electric potential as defined in Eq. 29-2 is a scalar because W and q_0 in that equation are scalars.

Both W_{AB} and $V_B - V_A$ in Eq. 29-1 are independent of the path followed in moving the test charge from point A to point B. If this were not so, point B would not have a unique electric potential (with respect to point A as a defined reference position) and the concept of potential would have limited usefulness.

We can easily prove that potential differences are path-independent for the special case shown in Fig. 29-1. This figure illustrates the case in which the two points A and B are in a field set up by a spherical charge q; the two points are further chosen, for simplicity, to lie along a radial line. Although our

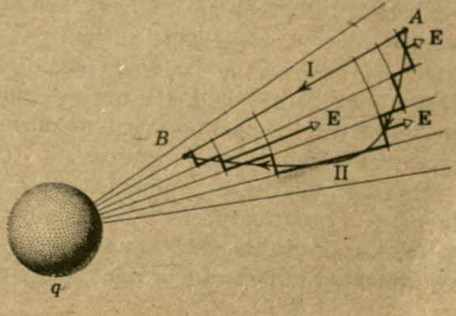


Fig. 29-1 A test charge q_0 is moved from A to B in the field of charge q along either of two paths. The open arrows show **E** at three points on path II.

path-independence proof applies only to this special case, it illustrates the general principles involved.

Point A in Fig. 29–1 may be taken as a defined reference point, and we imagine a positive test charge q_0 moved by an external agent from A to B. We consider two paths, path I being a radial line between A and B and path II being a completely arbitrary path between these two points. The open arrows on path II show the electric force per unit charge that would act at various points on a test charge q_0.

Path II may be approximated by a broken path made up of alternating elements of arc and of radius. Since these elements can be arbitrarily small, the broken path can be made arbitrarily close to the actual path. On path II the external agent does work *only along the radial segments* because along the arcs the force $\mathbf{F}$ and the displacement $d\mathbf{l}$ are at right angles, $\mathbf{F}\cdot d\mathbf{l}$ being zero in such cases. The sum of the work done on the radial segments that make up path II is the same as the work done on path I because each path has the same array of radial segments. Since path II is arbitrary, we have proved that the work done is the same for *all* paths connecting A and B. Although this proof holds only for the special case of Fig. 29–1, the potential difference is path-independent for *any* two points in *any* electrostatic field. We discussed path independence in Section 8–2 for the general class of *conservative forces;* electrostatic forces, like gravitational forces, are conservative.

The locus of points, all of which have the same electric potential, is called an *equipotential surface*. A family of equipotential surfaces, each surface corresponding to a different value of the potential, can be used to give a general description of the electric field in a certain region of space. We have seen earlier (Section 27–3) that electric lines of force can also be used for this purpose; in later sections (see, for example, Fig. 29–15) we explore the intimate connection between these two ways of describing the electric field.

No work is required to move a test charge between any two points on an equipotential surface. This follows from Eq. 29–1,

$$V_B - V_A = \frac{W_{AB}}{q_0},$$

because W_{AB} must be zero if $V_A = V_B$. This is true, because of the path independence of potential difference, even if the path connecting A and B does not lie entirely in the equipotential surface.

Figure 29–2 shows an arbitrary family of equipotential surfaces. The work to move a charge along paths I and II is zero because all these paths begin and end on the same equipotential surface. The work to move a charge along paths I′ and II′ is not zero but is the same for each path because the initial and the final potentials are identical; paths I′ and II′ connect the same pair of equipotential surfaces.

From symmetry, the equipotential surfaces for a spherical charge are a family of concentric spheres. For a uniform field they are a family of planes at right angles to the field. In all cases (including these two examples) the equipotential surfaces are at right angles to the lines of force and thus to $\mathbf{E}$

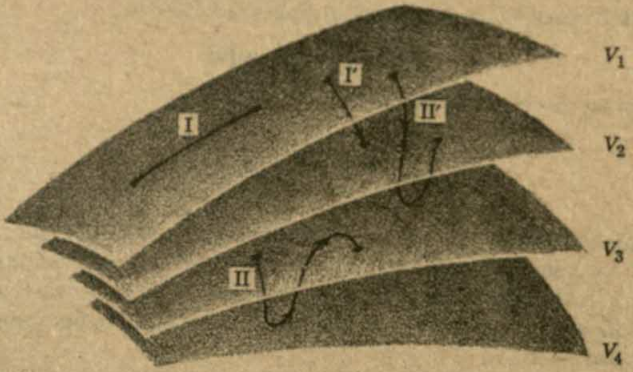


Fig. 29–2 Portions of four equipotential surfaces. The heavy lines show four paths along which a test charge is moved.

(see Fig. 29–15). If **E** were *not* at right angles to the equipotential surface, it would have a component lying in that surface. Then work would have to be done in moving a test charge about on the surface. Work cannot be done if the surface is an equipotential, so **E** must be at right angles to the surface.

There is a strong analogy between electrostatic forces and gravitational forces, based on the fact that their fundamental laws are inverse square laws (see Eqs. 26–3 and 16–1):

$$F_E = \frac{1}{4\pi\epsilon_0}\frac{q_1q_2}{r^2} \quad \text{and} \quad F_g = G\frac{m_1m_2}{r^2}.$$

Thus we can define the *gravitational potential* V_g (compare Eq. 29–2) from

$$V_g = \frac{W}{m},$$

where W is the work required to move a test body of mass m from infinity to the point in question. Gravitational equipotential surfaces can also be constructed; they prove to be everywhere at right angles to the gravitational field strength vector **g**. For a uniform gravitational field, such as that near the surface of the earth, these surfaces are horizontal planes. This correlates with the facts that (a) no net work is required to move a stone of mass m between two points with the same elevation and (b) the same net work is required to move a stone along any path starting on a given horizontal surface and ending on another.

29–2 Potential and Field Strength

Let A and B in Fig. 29–3 be two points in a uniform electric field **E**, set up by an arrangement of charges not shown, and let A be a distance d in the field direction from B. Assume that a positive test charge q_0 is moved, by an external agent and without acceleration, from A to B along the straight line connecting them.

The *electric* force on the charge is $q_0\mathbf{E}$ and points down. To move the charge in the way we have described we must counteract this force by applying an external force **F** of the same magnitude but directed upward. The work W done by the agent that supplies this force is

$$W_{AB} = Fd = q_0 Ed. \tag{29-3}$$

Substituting this into Eq. 29–1 yields

$$V_B - V_A = \frac{W_{AB}}{q_0} = Ed. \tag{29-4}$$

This equation shows the connection between potential difference and field strength for a simple special case. Note from this equation that another mks unit for **E** is the volt/meter. The student may wish to prove that a volt/meter is identical with a nt/coul; this latter unit is the one first presented for **E** in Section 27–2.

In Fig. 29–3 B has a higher potential than A. This is reasonable because an external agent would have to do positive work to push a positive test charge from A to B. Figure 29–3 could be used as it stands to illustrate the act of lifting a stone from A to B in the uniform gravitational field near the earth's surface.

What is the connection between V and **E** in the more general case in which the field is *not* uniform and in which the test body is moved along a path that is *not* straight, as in Fig. 29–4? The electric field exerts a force $q_0\mathbf{E}$ on the test charge, as shown. To keep the test charge from accelerating, an external agent must apply a force **F** chosen to be exactly equal to $-q_0\mathbf{E}$ for all positions of the test body.

If the external agent causes the test body to move through a displacement $d\mathbf{l}$ along the path from A to B, the element of work done by the external agent

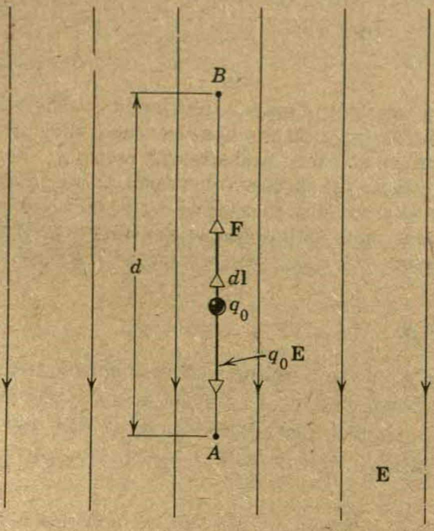


Fig. 29–3 A test charge q_0 is moved from A to B in a uniform electric field **E** by an external agent that exerts a force **F** on it.

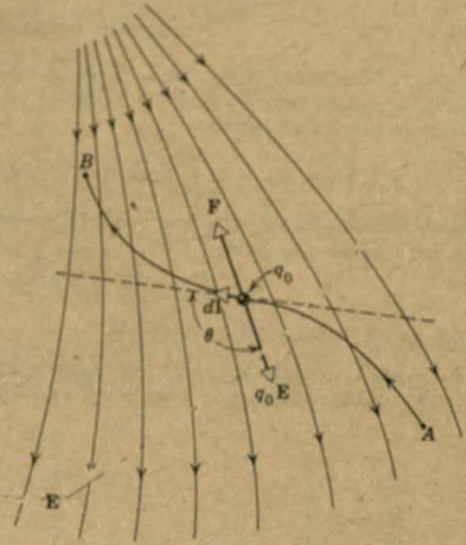


Fig. 29–4 A test charge q_0 is moved from A to B in a nonuniform electric field by an external agent that exerts a force $\mathbf{F}$ on it.

is $\mathbf{F}\cdot d\mathbf{l}$. To find the total work W_{AB} done by the external agent in moving the test charge from A to B, we add up (that is, integrate) the work contributions for all the infinitesimal segments into which the path is divided. This leads to

$$W_{AB} = \int_A^B \mathbf{F}\cdot d\mathbf{l} = -q_0\int_A^B \mathbf{E}\cdot d\mathbf{l}.$$

Such an integral is called a *line integral*. Note that we have substituted $-q_0\mathbf{E}$ for its equal, $\mathbf{F}$.

Substituting this expression for W_{AB} into Eq. 29–1 leads to

$$V_B - V_A = \frac{W_{AB}}{q_0} = -\int_A^B \mathbf{E}\cdot d\mathbf{l}. \tag{29–5}$$

If point A is taken to be infinitely distant and the potential V_A at infinity is taken to be zero, this equation gives the potential V at point B, or, dropping the subscript B,

$$V = -\int_\infty^B \mathbf{E}\cdot d\mathbf{l}. \tag{29–6}$$

These two equations allow us to calculate the potential difference between any two points (or the potential at any point) if $\mathbf{E}$ is known at various points in the field.

▶ **Example 1.** In Fig. 29-3 calculate $V_B - V_A$ using Eq. 29-5. Compare the result with that obtained by direct analysis of this special case (Eq. 29-4).

In moving the test charge the element of path $d\mathbf{l}$ always points in the direction of motion; this is upward in Fig. 29–3. The electric field **E** in this figure points down so that the angle θ between **E** and $d\mathbf{l}$ is 180°.

Equation 29–5 then becomes

$$V_B - V_A = -\int_A^B \mathbf{E}\cdot d\mathbf{l} = -\int_A^B E\cos 180°\, dl = \int_A^B E\, dl.$$

E is the constant for all parts of the path in this problem and can thus be taken outside the integral sign, giving

$$V_B - V_A = E\int_A^B dl = Ed,$$

which agrees with Eq. 29–4, as it must.

Example 2. In Fig. 29–5 let a test charge q_0 be moved without acceleration from A to B over the path shown. Compute the potential difference between A and B.

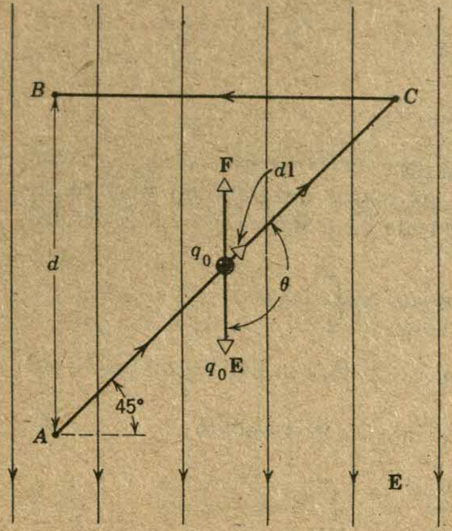


Fig. 29–5 Example 2. A test charge q_0 is moved along path ACB in a uniform electric field by an external agent.

For the path AC we have $\theta = 135°$ and, from Eq. 29–5,

$$V_C - V_A = -\int_A^C \mathbf{E}\cdot d\mathbf{l} = -\int_A^C E\cos 135°\, dl = \frac{E}{\sqrt{2}}\int_A^C dl.$$

The integral is the length of the line AC which is $\sqrt{2}d$. Thus

$$V_C - V_A = \frac{E}{\sqrt{2}}(\sqrt{2}d) = Ed.$$

Points B and C have the same potential because no work is done in moving a charge between them, **E** and $d\mathbf{l}$ being at right angles for all points on the line CB. In other words, B and C lie on the same equipotential surface at right angles to the lines of force. Thus

$$V_B - V_A = V_C - V_A = Ed.$$

This is the same value derived for a direct path connecting A and B, a result to be expected because the potential difference between two points is path independent. ◀

29-3 Potential Due to a Point Charge

Figure 29-6 shows two points A and B near an isolated point charge q. For simplicity we assume that A, B, and q lie on a straight line. Let us compute the potential difference between points A and B, assuming that a test charge q_0 is moved without acceleration along a radial line from A to B.

In Fig. 29-6 **E** points to the right and $d\mathbf{l}$, which is always in the direction of motion, points to the left. Therefore, in Eq. 29-5,

$$\mathbf{E}\cdot d\mathbf{l} = E \cos 180^\circ \, dl = -E \, dl.$$

However, as we move a distance dl to the left, we are moving in the direction of decreasing r because r is measured from q as an origin. Thus

$$dl = -dr.$$

Combining yields

$$\mathbf{E}\cdot d\mathbf{l} = E \, dr.$$

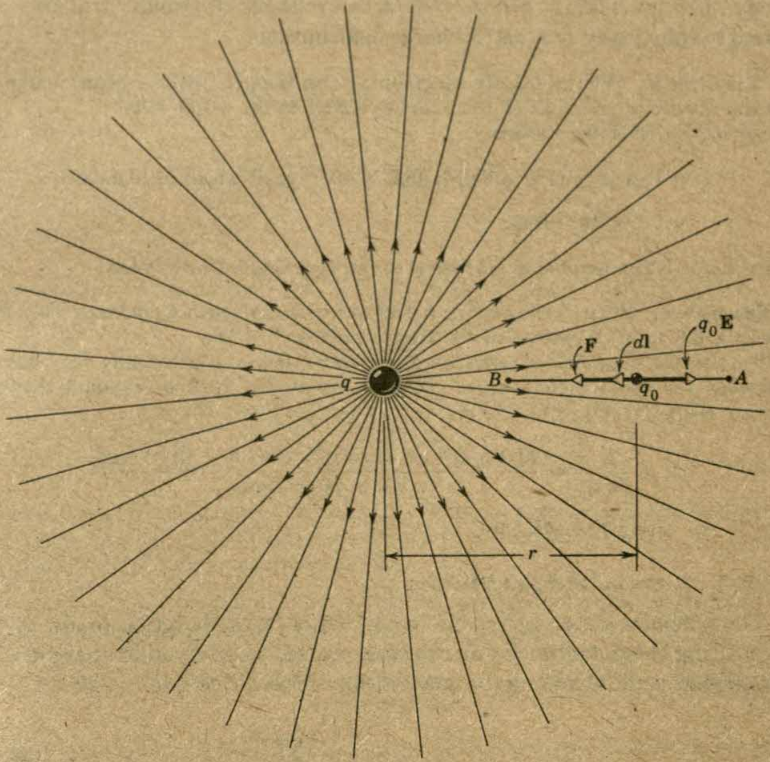


Fig. 29-6 A test charge q_0 is moved by an external agent from A to B in the field set up by a point charge q.

Substituting this into Eq. 29–5 gives

$$V_B - V_A = -\int_A^B \mathbf{E}\cdot d\mathbf{l} = -\int_{r_A}^{r_B} E\,dr.$$

Combining with Eq. 27–4,

$$E = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2}$$

leads to
$$V_B - V_A = -\frac{q}{4\pi\epsilon_0}\int_{r_A}^{r_B}\frac{dr}{r^2} = \frac{q}{4\pi\epsilon_0}\left(\frac{1}{r_B} - \frac{1}{r_A}\right). \tag{29–7}$$

Choosing reference position A to be at infinity (that is, letting $r_A \to \infty$), choosing $V_A = 0$ at this position, and dropping the subscript B leads to

$$V = \frac{1}{4\pi\epsilon_0}\frac{q}{r}. \tag{29–8}$$

This equation shows clearly that equipotential surfaces for an isolated point charge are spheres concentric with the point charge (see Fig. 29-15*a*). A study of the derivation will show that this relation also holds for points external to spherically symmetric charge distributions.

▶ **Example 3.** What must the magnitude of an isolated positive point charge be for the electric potential at 10 cm from the charge to be +100 volts?

Solving Eq. 29–8 for q yields

$$\begin{aligned} q = V4\pi\epsilon_0 r &= (100\text{ volts})(4\pi)(8.9\times10^{-12}\text{ coul}^2/\text{nt-m}^2)(0.10\text{ meter}) \\ &= 1.1\times10^{-9}\text{ coul.}\end{aligned}$$

This charge is comparable to charges that can be produced by friction.

Example 4. What is the electric potential at the surface of a gold nucleus? The radius is 6.6×10^{-15} meter and the atomic number $Z = 79$.

The nucleus, assumed spherically symmetrical, behaves electrically for external points as if it were a point charge. Thus we can use Eq. 29–8, or, recalling that the proton charge is 1.6×10^{-19} coul,

$$\begin{aligned} V = \frac{1}{4\pi\epsilon_0}\frac{q}{r} &= \frac{(9.0\times10^9\text{ nt-m}^2/\text{coul}^2)(79)(1.6\times10^{-19}\text{ coul})}{6.6\times10^{-15}\text{ meter}} \\ &= 1.7\times10^7\text{ volts.}\end{aligned}$$

◀

29–4 A Group of Point Charges

The potential at any point due to a group of point charges is found by (*a*) calculating the potential V_n due to each charge, as if the other charges were not present, and (*b*) adding the quantities so obtained, or (see Eq. 29–8)

$$V = \sum_n V_n = \frac{1}{4\pi\epsilon_0}\sum_n \frac{q_n}{r_n}, \tag{29–9}$$

where q_n is the value of the nth charge and r_n is the distance of this charge from the point in question. The sum used to calculate V is an *algebraic sum* and not a vector sum like the one used to calculate **E** for a group of point charges (see Eq. 27-5). Herein lies an important computational advantage of potential over electric field strength.

If the charge distribution is continuous, rather than being a collection of points, the sum in Eq. 29-9 must be replaced by an integral, or

$$V = \int dV = \frac{1}{4\pi\epsilon_0}\int \frac{dq}{r}. \tag{29-10}$$

where dq is a differential element of the charge distribution, r is its distance from the point at which V is to be calculated, and dV is the potential it establishes at that point.

▶ **Example 5.** What is the potential at the center of the square of Fig. 29-7? Assume that $q_1 = +1.0 \times 10^{-8}$ coul, $q_2 = -2.0 \times 10^{-8}$ coul, $q_3 = +3.0 \times 10^{-8}$ coul, $q_4 = +2.0 \times 10^{-8}$ coul, and $a = 1.0$ meter.

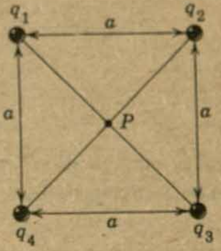


Fig. 29-7 Example 5

The distance r of each charge from P is $a/\sqrt{2}$ or 0.71 meter. From Eq. 29-9

$$V = \sum_n V_n = \frac{1}{4\pi\epsilon_0}\frac{q_1 + q_2 + q_3 + q_4}{r}$$

$$= \frac{(9.0 \times 10^9 \text{ nt-m}^2/\text{coul}^2)(1.0 - 2.0 + 3.0 + 2.0) \times 10^{-8}\text{ coul}}{0.71 \text{ meter}}$$

$$= 500 \text{ volts.}$$

Is the potential constant within the square? Does any point inside have a negative potential? Can you sketch roughly the intersection of the plane of Fig. 29-7 with the equipotential surface corresponding to zero volts?

Example 6. *A charged disk.* Find the electric potential for points on the axis of a uniformly charged circular disk whose surface charge density is σ (see Fig. 29-8).

Consider a charge element dq consisting of a flat circular strip of radius y and width dy. We have

$$dq = \sigma(2\pi y)(dy),$$

where $(2\pi y)(dy)$ is the area of the strip. All parts of this charge element are the same distance r' $(= \sqrt{y^2 + r^2})$ from axial point P so that their contribution dV to the

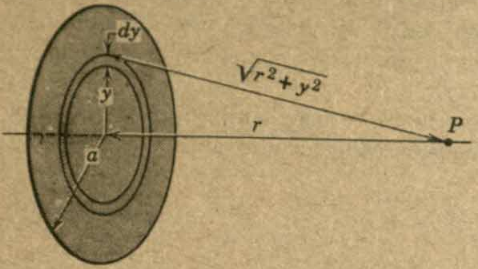


Fig. 29–8 Example 6. A point P on the axis of a uniformly charged circular disk of radius a.

electric potential at P is given by Eq. 29–8, or

$$dV = \frac{1}{4\pi\epsilon_0}\frac{dq}{r'} = \frac{1}{4\pi\epsilon_0}\frac{\sigma 2\pi y\,dy}{\sqrt{y^2+r^2}}.$$

The potential V is found by integrating over all the strips into which the disk can be divided (Eq. 29–10) or

$$V = \int dV = \frac{\sigma}{2\epsilon_0}\int_0^a (y^2+r^2)^{-1/2} y\,dy$$

$$= \frac{\sigma}{2\epsilon_0}(\sqrt{a^2+r^2} - r).$$

This general result is valid for all values of r. In the special case of $r \gg a$ the quantity $\sqrt{a^2+r^2}$ can be approximated as

$$\sqrt{a^2+r^2} = r\left(1+\frac{a^2}{r^2}\right)^{1/2} = r\left(1+\frac{1}{2}\frac{a^2}{r^2}+\cdots\right) \cong r+\frac{a^2}{2r},$$

in which the quantity in parentheses in the second member of this equation has been expanded by the binomial theorem (see Appendix *I*). This equation means that V becomes

$$V \cong \frac{\sigma}{2\epsilon_0}\left(r+\frac{a^2}{2r}-r\right) = \frac{\sigma\pi a^2}{4\pi\epsilon_0 r} = \frac{1}{4\pi\epsilon_0}\frac{q}{r},$$

where q $(= \sigma\pi a^2)$ is the total charge on the disk. This limiting result is expected because the disk behaves like a point charge for $r \gg a$. ◀

29–5 Potential Due to a Dipole

Two equal charges, q, of opposite sign, separated by a distance $2a$, constitute an electric dipole; see Example 3, Chapter 27. The electric dipole moment **p** has the magnitude $2aq$ and points from the negative charge to the positive charge. Here we derive an expression for the electric potential V at any point of space due to a dipole, provided only that the point is not too close to the dipole.

A point P is specified by giving the quantities r and θ in Fig. 29–9. From symmetry, it is clear that the potential will not change as point P rotates about the z axis, r and θ being fixed. Thus we need only find $V(r,\theta)$ for any plane containing this axis; the plane of Fig. 29–9 is such a plane. Applying

Eq. 29-9 gives

$$V = \sum_n V_n = V_1 + V_2 = \frac{1}{4\pi\epsilon_0}\left(\frac{q}{r_1} - \frac{q}{r_2}\right) = \frac{q}{4\pi\epsilon_0}\frac{r_2 - r_1}{r_1 r_2},$$

which is an exact relationship.

We now limit consideration to points such that $r \gg 2a$. These approximate relations then follow from Fig. 29-9:

$$r_2 - r_1 \cong 2a\cos\theta \qquad \text{and} \qquad r_1 r_2 \cong r^2,$$

and the potential reduces to

$$V = \frac{q}{4\pi\epsilon_0}\frac{2a\cos\theta}{r^2} = \frac{1}{4\pi\epsilon_0}\frac{p\cos\theta}{r^2}, \tag{29-11}$$

in which p $(= 2aq)$ is the dipole moment. Note that V vanishes everywhere in the equatorial plane ($\theta = 90°$). This reflects the fact that it takes no work to bring a test charge in from infinity along the perpendicular bisector of the dipole. For a given radius, V has its greatest positive value for $\theta = 0$ and its greatest negative value for $\theta = 180°$. Note that the potential does not depend separately on q and $2a$ but only on their product p.

It is convenient to call *any* assembly of charges, for which V at distant points is given by Eq. 29-11, an *electric dipole*. Two point charges separated by a small distance behave this way, as we have just proved. However, other charge configurations also obey Eq. 29-11. Suppose that by measurement at points outside an imaginary box (Fig. 29-10) we find a pattern of lines of force that can be described quantitatively by Eq. 29-11. We then declare that the object inside the box is an *electric dipole*, that its axis is the line zz', and that its dipole moment **p** points vertically upward.

Many molecules have electric dipole moments. That for H_2O in its vapor state is 6.1×10^{-30} coul-m. Figure 29-11 is a representation of this molecule, showing the three nuclei and the surrounding electron cloud. The dipole moment **p** is represented

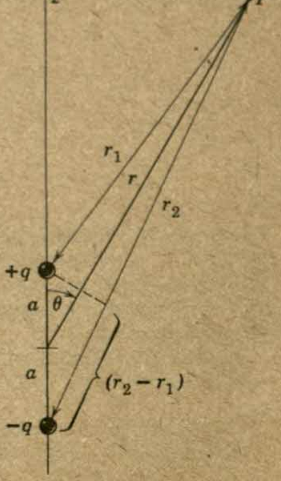


Fig. 29-9 A point P in the field of an electric dipole.

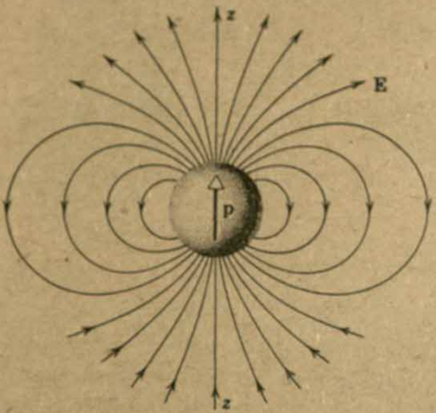


Fig. 29–10 If an object inside the spherical box sets up the electric field shown (described quantitatively by Eq. 29–11), it is an electric dipole.

by the arrow on the axis of symmetry of the molecule. In this molecule the effective center of positive charge does not coincide with the effective center of negative charge. It is precisely because of this separation that the dipole moment exists.

Atoms, and many molecules, do *not* have permanent dipole moments. However, dipole moments may be induced by placing any atom or molecule in an external electric field. The action of the field (Fig. 29–12) is to separate the centers of positive and of negative charge. We say that the atom becomes *polarized* and acquires an *induced electric dipole moment*. Induced dipole moments disappear when the electric field is removed.

Electric dipoles are important in situations other than atomic and molecular ones. Radio and radar antennas are often in the form of a metal wire or rod in which electrons surge back and forth periodically. At a certain time one end of the wire or rod will be negative and the other end positive. Half a cycle later the polarity of the ends is exactly reversed. This is an *oscillating* electric dipole. It is so named because its dipole moment changes in a periodic way with time.

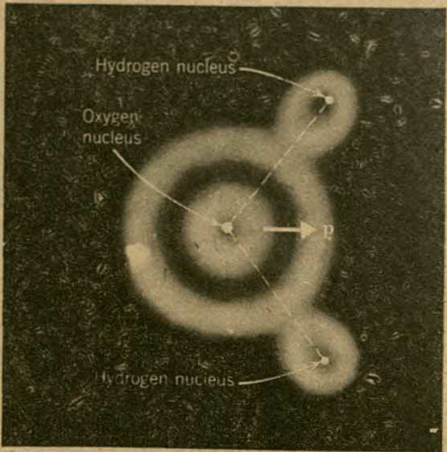


Fig. 29–11 A schematic representation of a water molecule, showing the three nuclei, the electron cloud, and the orientation of the dipole moment.

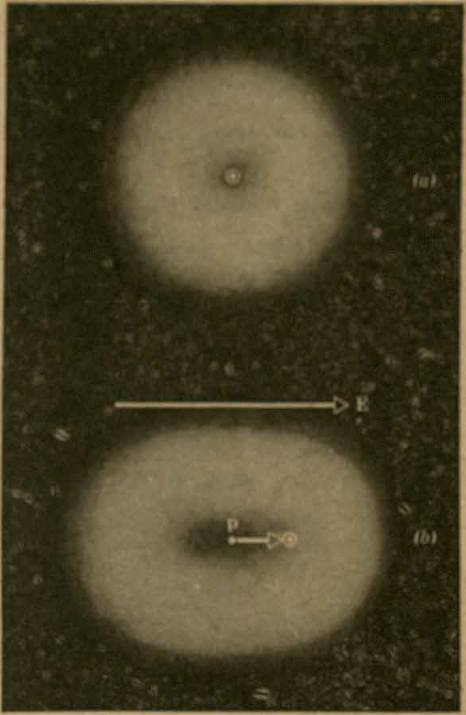


Fig. 29-12 (*a*) An atom, showing the nucleus and the electron cloud. The center of negative charge coincides with the center of positive charge, that is, with the nucleus. (*b*) If an external field **E** is applied, the electron cloud is distorted so that the center of negative charge, marked by the dot, and the center of positive charge no longer coincide. An electric dipole appears.

▶ **Example 7.** *An electric quadrupole.* An electric quadrupole, of which Fig. 27-21 is an example, consists of two electric dipoles so arranged that they almost, but not quite, cancel each other in their electric effects at distant points.* Calculate $V(r)$ for points on the axis of this quadrupole.

Applying Eq. 29-9 to Fig. 27-21 yields

$$V = \sum_n V_n = \frac{1}{4\pi\epsilon_0}\left(\frac{q}{r-a} - \frac{2q}{r} + \frac{q}{r+a}\right)$$

$$= \frac{q}{4\pi\epsilon_0}\frac{2a^2}{(r-a)(r)(r+a)}.$$

Assuming that $r \gg a$ allows us to put $a = 0$ in the sum and difference terms in the denominator, yielding

$$V = \frac{1}{4\pi\epsilon_0}\frac{Q}{r^3},$$

where Q ($= 2qa^2$) is the *electric quadrupole moment* of the charge assembly of Fig. 27-21. Note that V varies (*a*) as $1/r$ for a point charge (see Eq. 29-8), (*b*) as $1/r^2$ for a dipole (see Eq. 29-11), and (*c*) as $1/r^3$ for a quadrupole.

Note too that (*a*) a dipole is two equal and opposite charges that do not coincide in space so that their electric effects at distant points do not quite cancel, and (*b*) a quadrupole is two equal and opposite dipoles that do not coincide in space so that their electric effects at distant points again do not quite cancel. This pattern can be extended to define higher orders of charge distribution such as *octupoles*.

---

* See Problem 18, Chapter 27.

The potential at points at distances from an *arbitrary* charge distribution (continuous or discrete) that are large compared with the size of the distribution can always be written as the sum of separate potential distributions due to (*a*) a single charge—sometimes, in this context, called a *monopole*—(*b*) a dipole, (*c*) a quadrupole, etc. This process is called an *expansion in multipoles* and is a very useful technique in problem solving. ◀

29-6 Electric Potential Energy *

If we raise a stone from the earth's surface, the work that we do against the earth's gravitational attraction is stored as *potential energy* in the system earth + stone. If we release the stone, the stored potential energy changes steadily into kinetic energy as the stone drops. After the stone comes to rest on the earth, this kinetic energy, equal in magnitude just before the time of contact to the originally stored potential energy, is transformed into heat energy in the system earth + stone.

A similar situation exists in electrostatics. Consider two charges q_1 and q_2 a distance r apart, as in Fig. 29-13. If we increase the separation between them, an external agent must do work that will be positive if the charges are opposite in sign and negative otherwise. The energy represented by this work can be thought of as stored in the system $q_1 + q_2$ as *electric potential energy*. This energy, like all varieties of potential energy, can be transformed into other forms. If q_1 and q_2, for example, are charges of opposite sign and we release them, they will accelerate toward each other, transforming the stored potential energy into kinetic energy of the accelerating masses. The analogy to the earth + stone system is exact, save for the fact that electric forces may be either attractive or repulsive whereas gravitational forces are always attractive.

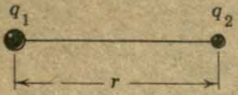


Fig. 29-13

We define the electric potential energy of a system of point charges as the work required to assemble this system of charges by bringing them in from an infinite distance. We assume that the charges are all at rest when they are infinitely separated, that is, they have no initial kinetic energy.

In Fig. 29-13 let us imagine q_2 removed to infinity and at rest. The

---

* In mechanics the concept of *potential energy* (of compressed springs, falling masses, etc.) is more commonly used than the concept of *potential*. In electrostatics the reverse is true, electric potential being perhaps a more common concept than electric potential energy. In what follows the student must be careful not to confuse these quite different quantities, potential and potential energy.

electric potential at the original site of q_2, caused by q_1, is given by Eq. 29-8, or

$$V = \frac{1}{4\pi\epsilon_0}\frac{q_1}{r}.$$

If q_2 is moved in from infinity to the original distance r, the work required is, from the definition of electric potential, that is, from Eq. 29-2,

$$W = Vq_2. \tag{29-12}$$

Combining these two equations and recalling that this work W is precisely the *electric potential energy* U of the system $q_1 + q_2$ yields

$$U\ (= W) = \frac{1}{4\pi\epsilon_0}\frac{q_1q_2}{r_{12}}. \tag{29-13}$$

The subscript of r emphasizes that the distance involved is that between the point charges q_1 and q_2.

For systems containing more than two charges the procedure is to compute the potential energy for every pair of charges separately and to add the results algebraically. This procedure rests on a physical picture in which (*a*) charge q_1 is brought into position, (*b*) q_2 is brought from infinity to its position near q_1, (*c*) q_3 is brought from infinity to its position near q_1 and q_2, etc.

The potential energy of continuous charge distributions (an ellipsoid of charge, for example) can be found by dividing the distribution into infinitesimal elements dq, treating each such element as a point charge, and using the procedures of the preceding paragraph, with the summation process replaced by an integration. We have not considered such problems in this text.

▶ **Example 8.** Two protons in a nucleus of U^{238} are 6.0×10^{-15} meter apart. What is their mutual electric potential energy?

From Eq. 29-13

$$U = \frac{1}{4\pi\epsilon_0}\frac{q_1q_2}{r} = \frac{(9.0 \times 10^9 \text{ nt-m}^2/\text{coul}^2)(1.6 \times 10^{-19} \text{ coul})^2}{6.0 \times 10^{-15} \text{ meter}}$$

$$= 3.8 \times 10^{-14} \text{ joule} = 2.4 \times 10^5 \text{ ev}.$$

Example 9. Three charges are arranged as in Fig. 29-14. What is their mutual potential energy? Assume that $q = 1.0 \times 10^{-7}$ coul and $a = 10$ cm.

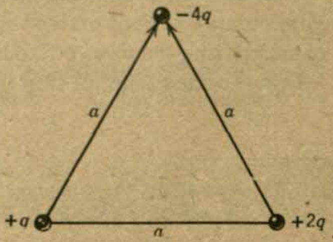


Fig. 29-14 Example 9. Three charges are fixed rigidly, as shown, by external forces.

The total energy of the configuration is the sum of the energies of each pair of particles. From Eq. 29–13,

$$
\begin{aligned}
U &= U_{12} + U_{13} + U_{23} \\
&= \frac{1}{4\pi\epsilon_0}\left[\frac{(+q)(-4q)}{a} + \frac{(+q)(+2q)}{a} + \frac{(-4q)(+2q)}{a}\right] \\
&= -\frac{10}{4\pi\epsilon_0}\frac{q^2}{a} \\
&= -\frac{(9.0 + 10^9\ \text{nt-m}^2/\text{coul}^2)(10)(1.0 \times 10^{-7}\ \text{coul})^2}{0.10\ \text{meter}} = -9.0 \times 10^{-3}\ \text{joule}.
\end{aligned}
$$

The fact that the total energy is negative means that negative work would have to be done to assemble this structure, starting with the three charges separated and at rest at infinity. Expressed otherwise, 9.0×10^{-3} joule of work must be done to dismantle this structure, removing the charges to an infinite separation from one another.

When, as is common practice, infinity is taken as the zero of electric potential, a positive potential energy (as in Example 8) corresponds to repulsive electric forces and a negative potential energy (as in this example) to attractive electric forces. If the protons in Example 8 were not held in place by attractive (nonelectrical) nuclear forces, they would move away from each other. If the three particles in this example were released from their fixed positions, in which they are held by external forces, they would move toward each other. ◀

29–7 Calculation of E from *V*

We have stated that V and **E** are equivalent descriptions and have determined (Eq. 29–6) how to calculate V from **E**. Let us now consider how to calculate **E** if V is known throughout a certain region.

This problem has already been solved graphically. If **E** is known at every point in space, the lines of force can be drawn; then a family of equipotentials can be sketched in by drawing surfaces at right angles. These equipotentials describe the behavior of V. Conversely, if V is given as a function of position, a set of equipotential surfaces can be drawn. The lines of force can then be found by drawing lines at right angles, thus describing the behavior of **E**. It is the mathematical equivalent of this second graphical process that we seek here. Figure 29–15 shows some examples of lines of force and of the corresponding equipotential surfaces.

Figure 29–16 shows the intersection with the plane of the figure of a family of equipotential surfaces. The figure shows that **E** at a typical point P is at right angles to the equipotential surface through P, as it must be.

Let us move a test charge q_0 from P along the path marked $\Delta\mathbf{l}$ to the equipotential surface marked $V + \Delta V$. The work that must be done by the agent exerting the force **F** (see Eq. 29–1) is $q_0\Delta V$.

From another point of view we can calculate the work from *

$$\Delta W = \mathbf{F}\cdot\Delta\mathbf{l},$$

---

* We assume that the equipotentials are so close together that **F** is constant for all parts of the path $\Delta\mathbf{l}$. In the limit of a differential path ($d\mathbf{l}$) there will be no difficulty.

where **F** is the force that must be exerted on the charge to overcome exactly the electrical force $q_0\mathbf{E}$. Since **F** and $q_0\mathbf{E}$ have opposite signs and are equal in magnitude,

$$\Delta W = -q_0\mathbf{E}\cdot\Delta\mathbf{l} = -q_0E\cos(\pi-\theta)\,\Delta l = q_0E\cos\theta\,\Delta l.$$

These two expressions for the work must be equal, which gives

$$q_0\,\Delta V = q_0E\cos\theta\,\Delta l$$

or

$$E\cos\theta = \frac{\Delta V}{\Delta l}. \tag{29-14}$$

Now $E\cos\theta$ is the component of **E** in the direction $-\mathbf{l}$ in Fig. 29-16; the quantity $-E\cos\theta$, which we call E_l, would then be the component of **E** in the $+\mathbf{l}$ direction. In the differential limit Eq. 29-14 can then be written as

$$E_l = -\frac{dV}{dl}. \tag{29-15}$$

In words, this equation says: If we travel through an electric field along a straight line and measure V as we go, the rate of change of V with distance that we observe, when changed in sign, is the component of **E** in that direction. The minus sign implies that **E** points in the direction of decreasing V, as in Fig. 29-16. It is particularly clear from Eq. 29-15 that appropriate units for **E** are the volts/meter.

There will be one direction **l** for which the quantity dV/dl is a maximum. From Eq. 29-15, E_l will also be a maximum for this direction and will in fact be E itself. Thus

$$E = -\left(\frac{dV}{dl}\right)_{\max} \tag{29-16}$$

The maximum value of dV/dl at a given point is called the *potential gradient* at that point. The direction **l** for which dV/dl has its maximum value is always at right angles to the equipotential surface, corresponding to $\theta = 0$ in Fig. 29-16.

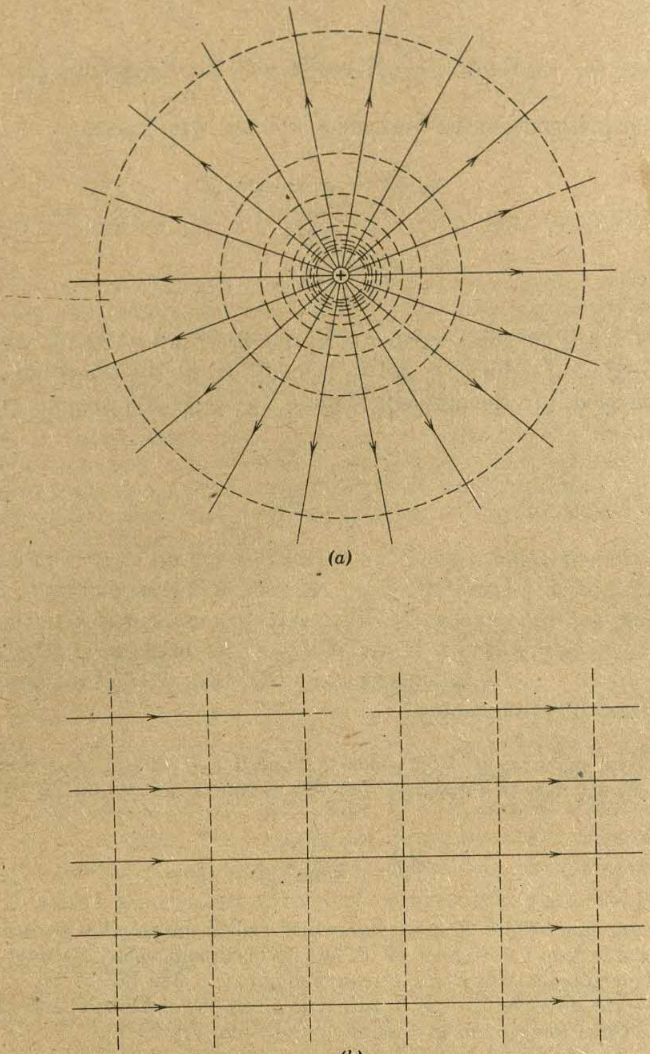
If we take the direction **l** to be, in turn, the directions of the x, y, and z axes, we can find the three components of **E** at any point, from Eq. 29-15.

$$E_x = -\frac{\partial V}{\partial x};\quad E_y = -\frac{\partial V}{\partial y};\quad E_z = -\frac{\partial V}{\partial z}. \tag{29-17}$$

Thus if V is known for all points of space, that is, if the function $V(x, y, z)$ is known the components of **E**, and thus **E** itself, can be found by taking derivatives.*

---

[*] The symbol $\partial V/\partial x$ is a *partial derivative*. It implies that in taking this derivative of the function $V(x, y, z)$ the quantity x is to be viewed as a variable and y and z are to be regarded as constants. Similar considerations hold for $\partial V/\partial y$ and $\partial V/\partial z$.



(a)

(b)

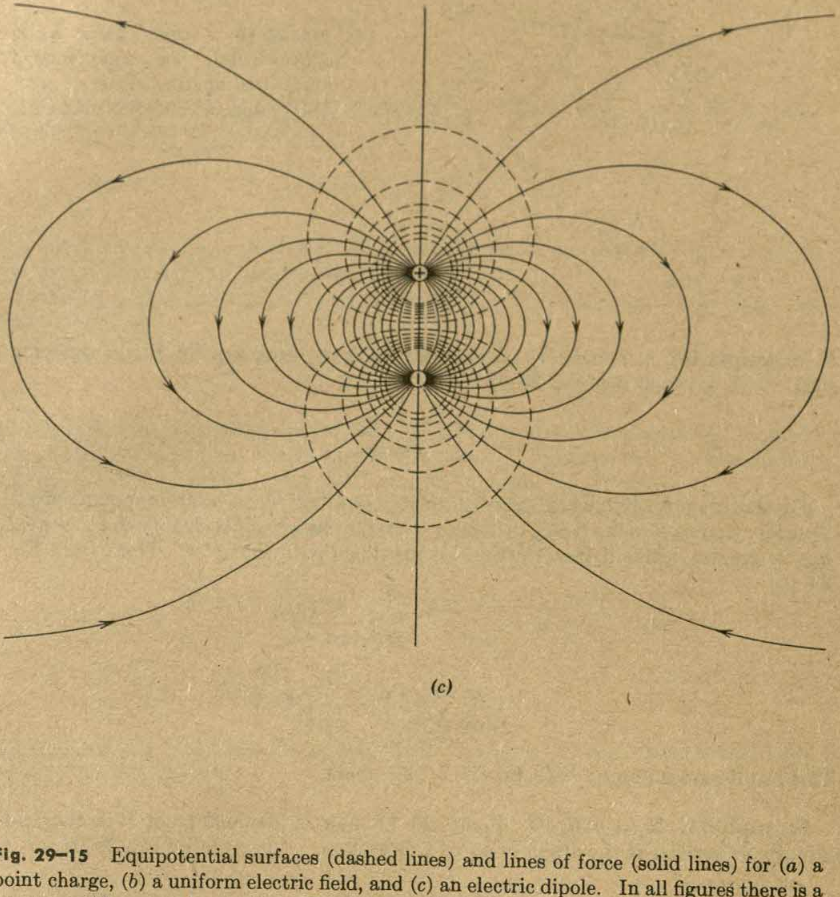


Fig. 29–15 Equipotential surfaces (dashed lines) and lines of force (solid lines) for (*a*) a point charge, (*b*) a uniform electric field, and (*c*) an electric dipole. In all figures there is a constant difference of potential ΔV between adjacent equipotential surfaces. Thus from Eq. 29–14, written for the case of $\theta = 180°$ as $\Delta l = -\Delta V/E$, the surfaces will be relatively close together where E is relatively large and relatively far apart where E is small. Similarly (see Section 27–3) the lines of force are relatively close together where E is large and far apart where E is small. See discussion and figures of Section 18–7 for other examples.

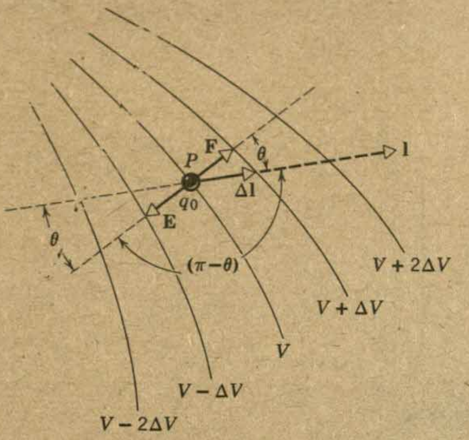


Fig. 29–16 A test charge q_0 is moved from one equipotential surface to another along an arbitrarily selected direction marked **l**.

▶ **Example 10.** Calculate $E(r)$ for a point charge q, using Eq. 29–16 and assuming that $V(r)$ is given as (see Eq. 29–8)

$$V = \frac{1}{4\pi\epsilon_0}\frac{q}{r}.$$

From symmetry, **E** must be directed radially outward for a (positive) point charge. Consider a point P in the field a distance r from the charge. It is clear that dV/dl at P has its greatest value if the direction l is identified with that of r. Thus, from Eq. 29–16,

$$E = -\frac{dV}{dr} = -\frac{d}{dr}\left(\frac{1}{4\pi\epsilon_0}\frac{q}{r}\right)$$

$$= -\frac{q}{4\pi\epsilon_0}\frac{d}{dr}\left(\frac{1}{r}\right) = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2}.$$

This result agrees exactly with Eq. 27–4, as it must.

Example 11. **E** *for a dipole.* Figure 29–17 shows a (distant) point P in the field of a dipole located at the origin of an xy-axis system. V is given by Eq. 29–11, or

$$V = \frac{1}{4\pi\epsilon_0}\frac{p\cos\theta}{r^2}.$$

Calculate **E** as a function of position.

From symmetry, **E**, for points in the plane of Fig. 29–17, lies in this plane. Thus it can be expressed in terms of its components E_x and E_y. Let us first express the potential function in rectangular coordinates rather than polar coordinates, making use of

$$r = (x^2 + y^2)^{1/2} \qquad \text{and} \qquad \cos\theta = \frac{y}{(x^2 + y^2)^{1/2}}.$$

The result is

$$V = \frac{p}{4\pi\epsilon_0}\frac{y}{(x^2 + y^2)^{3/2}}.$$

We find E_y from Eq. 29-17, recalling that x is to be treated as a constant in this calculation:

$$E_y = -\frac{\partial V}{\partial y} = -\frac{p}{4\pi\epsilon_0}\frac{(x^2+y^2)^{3/2} - y\frac{3}{2}(x^2+y^2)^{1/2}(2y)}{(x^2+y^2)^3}$$

$$= -\frac{p}{4\pi\epsilon_0}\frac{x^2-2y^2}{(x^2+y^2)^{5/2}}.$$

Note that putting $x = 0$ describes points along the dipole axis (that is, the y axis), and the expression for E_y reduces to

$$E_y = \frac{2p}{4\pi\epsilon_0}\frac{1}{y^3}.$$

This result agrees exactly with that found in Chapter 27 (see Problem 10), for, from symmetry, E_x equals zero on the dipole axis.

Putting $y = 0$ in the expression for E_y describes points in the median plane of the dipole and yields

$$E_y = -\frac{p}{4\pi\epsilon_0}\frac{1}{x^3},$$

which agrees exactly with the result found in Chapter 27 (see Example 3), for, again from symmetry, E_x equals zero in the median plane. The minus sign in this equation indicates that **E** points in the negative y direction (see Fig. 29-10).

The component E_x is also found from Eq. 29-17, recalling that y is to be taken as a constant during this calculation:

$$E_x = -\frac{\partial V}{\partial x} = -\frac{py}{4\pi\epsilon_0}(-\tfrac{3}{2})(x^2+y^2)^{-5/2}(2x)$$

$$= \frac{3p}{4\pi\epsilon_0}\frac{xy}{(x^2+y^2)^{5/2}}.$$

As expected, E_x vanishes both on the dipole axis ($x = 0$) and in the median plane ($y = 0$); see Fig. 29-10. ◀

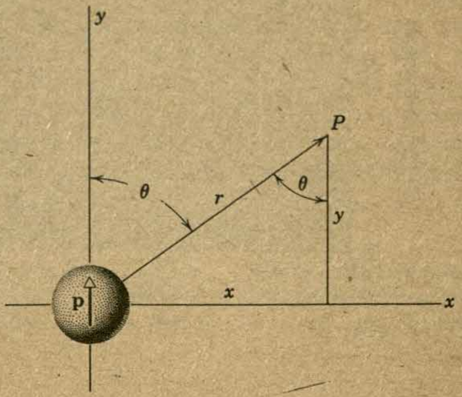


Fig. 29-17 Showing a point P in the field of an electric dipole **p**.

29-8 An Insulated Conductor

We proved in Section 28-4, using Gauss's law, that after a steady state is reached an excess charge q placed on an insulated conductor will move to its outer surface. We now assert that this charge q will distribute itself on this surface so that all points of the conductor, including *those on the surface and those inside*, have the same potential.

Consider any two points A and B in or on the conductor. If they were not at the same potential, the charge carriers in the conductor near the point of lower potential would tend to move toward the point of higher potential. We have assumed, however, that a steady-state situation, in which such currents do not exist, has been reached; thus all points, both on the surface and inside

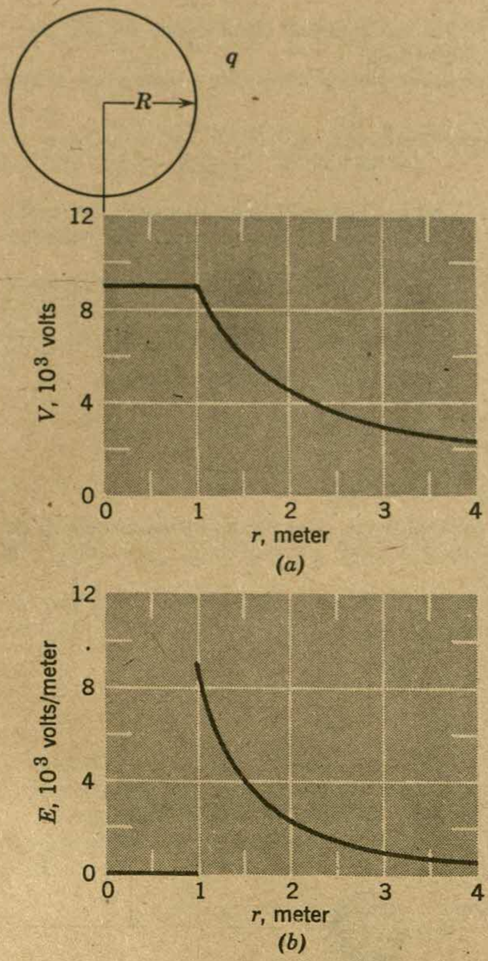


Fig. 29-18 (*a*) The potential and (*b*) the electric field strength, for points near a conducting spherical shell of 1.0 meter radius carrying a charge of $+1.0 \times 10^{-6}$ coul.

it, must have the same potential. Since the surface of the conductor is an equipotential surface, **E** for points on the surface must be at right angles to the surface.

We saw in Section 28-4 that a charge placed on an insulated conductor will spread over the surface until **E** equals zero for all points inside. We now have an alternative way of saying the same thing; the charge will move until all points of the conductor (surface points and interior points) are brought to the same potential, for if V is constant in the conductor then **E** is zero everywhere in the conductor ($E_l = -dV/dl$).

Figure 29-18*a* is a plot of potential against radial distance for an isolated spherical conducting shell of 1.0-meter radius carrying a positive charge of 1.0×10^{-6} coul. For points outside the shell $V(r)$ can be calculated from Eq. 29-8 because the charge q behaves, for such points, as if it were concentrated at the center of the sphere. Equation 29-8 is correct right up to the surface of the shell. Now let us push the test charge through the surface, assuming that there is a small hole, and into the interior. No extra work is needed because no electrical forces act on the test charge once it is inside the shell. Thus the potential everywhere inside is the same as that on the surface, as Fig. 29-18*a* shows.

Figure 29-18*b* shows the electric field strength for this same sphere. Note that E equals zero inside. The lower of these curves can be derived from the upper by differentiation with respect to r, using Eq. 29-16; the upper can be derived from the lower by integration with respect to r, using Eq. 29-6.

Figure 29-18 holds without change if the conductor is a solid sphere rather than a spherical shell as we have assumed. It is constructive to compare Fig. 29-18*b* (conducting shell or sphere) with Fig. 28-9, which holds for a *nonconducting* sphere. The student should try to understand the difference between these two figures, bearing in mind that in the first the charge lies on the surface whereas in the second it was assumed to have been spread uniformly throughout the volume of the sphere.

Finally we note that, as a general rule, the charge density tends to be high on isolated conducting surfaces whose radii of curvature are small, and conversely. For example, the charge density tends to be relatively high on sharp points and relatively low on plane regions on a conducting surface. The electric field strength E at points immediately above a charged surface is proportional to the charge density σ so that E may also reach very high values near sharp points. Glow discharges from sharp points during thunderstorms are a familiar example. The lightning rod acts in this way to neutralize charged clouds and thus prevent lightning strokes.

We can examine the qualitative relationship between σ and the curvature of the surface in a particular case by considering two spheres of different radii connected by a very long fine wire; see Fig. 29-19. Suppose that this entire assembly is raised

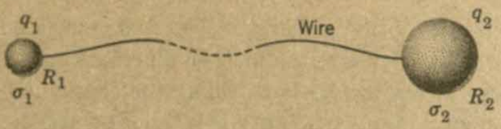


Fig. 29–19 Two spheres connected by a long fine wire.

to some arbitrary potential V. The (equal) potentials of the two spheres are, from Eq. 29–8,*

$$V = \frac{1}{4\pi\epsilon_0}\frac{q_1}{R_1} = \frac{1}{4\pi\epsilon_0}\frac{q_2}{R_2},$$

which yields

$$\frac{q_1}{q_2} = \frac{R_1}{R_2}, \tag{29–18}$$

where q_1 is the charge on the sphere of radius R_1 and q_2 is the charge on the sphere of radius R_2.

The *surface charge densities* for each sphere are given by

$$\sigma_1 = \frac{q_1}{4\pi R_1^2} \quad \text{and} \quad \sigma_2 = \frac{q_2}{4\pi R_2^2}.$$

Dividing gives

$$\frac{\sigma_1}{\sigma_2} = \frac{q_1}{q_2}\frac{R_2^2}{R_1^2}.$$

Combining with Eq. 29–18 yields

$$\frac{\sigma_1}{\sigma_2} = \frac{R_2}{R_1},$$

which is consistent with our qualitative statement above. Note that the larger sphere has the larger total charge but the smaller charge density.

The fact that σ, and thus **E**, can become very large near sharp points is important in the design of high-voltage equipment. *Corona discharge* can result from such points if the conducting object is raised to high potential and surrounded by air. Normally air is thought of as a nonconductor. However, it contains a small number of ions produced, for example, by the cosmic rays. A positively charged conductor will attract negative ions from the surrounding air and thus will slowly neutralize itself.

---

* Equation 29–8 holds only for an *isolated* point charge or spherically symmetric charge distribution. The spheres must be assumed to be so far apart that the charge on either one has a negligible effect on the distribution of charge on the other.

If the charged conductor has sharp points, the value of E in the air near the points can be very high. If the value is high enough, the ions, as they are drawn toward the conductor, will receive such large accelerations that, by collision with air molecules, they will produce vast additional numbers of ions. The air is thus made much more conducting, and the discharge of the conductor by this corona discharge may be very rapid indeed. The air surrounding sharp conducting points may even glow visibly because of light emitted from the air molecules during these collisions.

29-9 The Electrostatic Generator

The electrostatic generator was conceived by Lord Kelvin in 1890 and put into useful practice in essentially its modern form by R. J. Van de Graaff in 1931. It is a device for producing electric potential differences of the order of several millions of volts. Its chief application in physics is the use of this potential difference to accelerate charged particles to high energies. Beams of energetic particles made in this way can be used in many different "atom-smashing" experiments. The technique is to let a charged particle "fall" through a potential difference V, gaining kinetic energy as it does so.

Let a particle of (positive) charge q move in a vacuum under the influence of an electric field from one position A to another position B whose electric potential is lower by V. The electric potential energy of the system is reduced by qV because this is the work that an external agent would have to do to restore the system to its original condition. This decrease in potential energy appears as kinetic energy of the particle, or

$$K = qV. \qquad (29\text{-}19)$$

K is in joules if q is in coulombs and V in volts. If the particle is an electron or a proton, q will be the quantum of charge e.

If we adopt the quantum of charge e as a unit in place of the coulomb, we arrive at another unit for energy, the *electron volt*, which is used extensively in atomic and nuclear physics. By substituting into Eq. 29-19,

$$\begin{aligned} 1 \text{ electron volt} &= (1 \text{ quantum of charge})(1 \text{ volt}) \\ &= (1.60 \times 10^{-19} \text{ coul})(1.00 \text{ volt}) \\ &= 1.60 \times 10^{-19} \text{ joule.} \end{aligned}$$

The electron volt can be used interchangeably with any other energy unit. Thus a 10-gm object moving at 1000 cm/sec can be said to have a kinetic energy of 3.1×10^{18} ev. Most physicists would prefer to express this result as 0.50 joule, the electron volt being inconveniently small. In atomic and nuclear problems, however, the electron volt (ev) and its multiples the Mev ($= 10^6$ ev), the Bev ($= 10^9$ ev) and the Gev ($= 10^{12}$ ev) are the usual units of choice.

▶ **Example 12.** *The electrostatic generator.* Figure 29–20, which illustrates the basic operating principle of the electrostatic generator, shows a small sphere of radius r placed inside a large spherical shell of radius R. The two spheres carry charges q and Q, respectively. Calculate their potential difference.

The potential of the large sphere is caused in part by its own charge and in part because it lies in the field set up by the charge q on the small sphere. From Eq. 29–8,

$$V_R = \frac{1}{4\pi\epsilon_0}\left(\frac{Q}{R} + \frac{q}{R}\right).$$

The potential of the small sphere is caused in part by its own charge and in part because it is inside the large sphere; see Fig. 29–18*a*. From Eq. 29–8,

$$V_r = \frac{1}{4\pi\epsilon_0}\left(\frac{q}{r} + \frac{Q}{R}\right).$$

The potential difference is

$$V_r - V_R = \frac{q}{4\pi\epsilon_0}\left(\frac{1}{r} - \frac{1}{R}\right).$$

Thus, assuming q is positive, the inner sphere will always be higher in potential than the outer sphere. If the spheres are connected by a fine wire, the charge q will flow *entirely* to the outer sphere, regardless of the charge Q that may already be present.

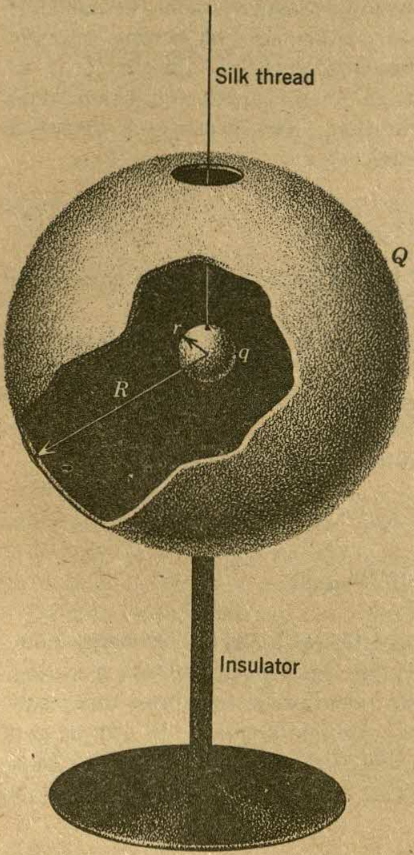


Fig. 29–20 Example 12. A small charged sphere of radius r is suspended inside a charged spherical shell whose outer surface has radius R.

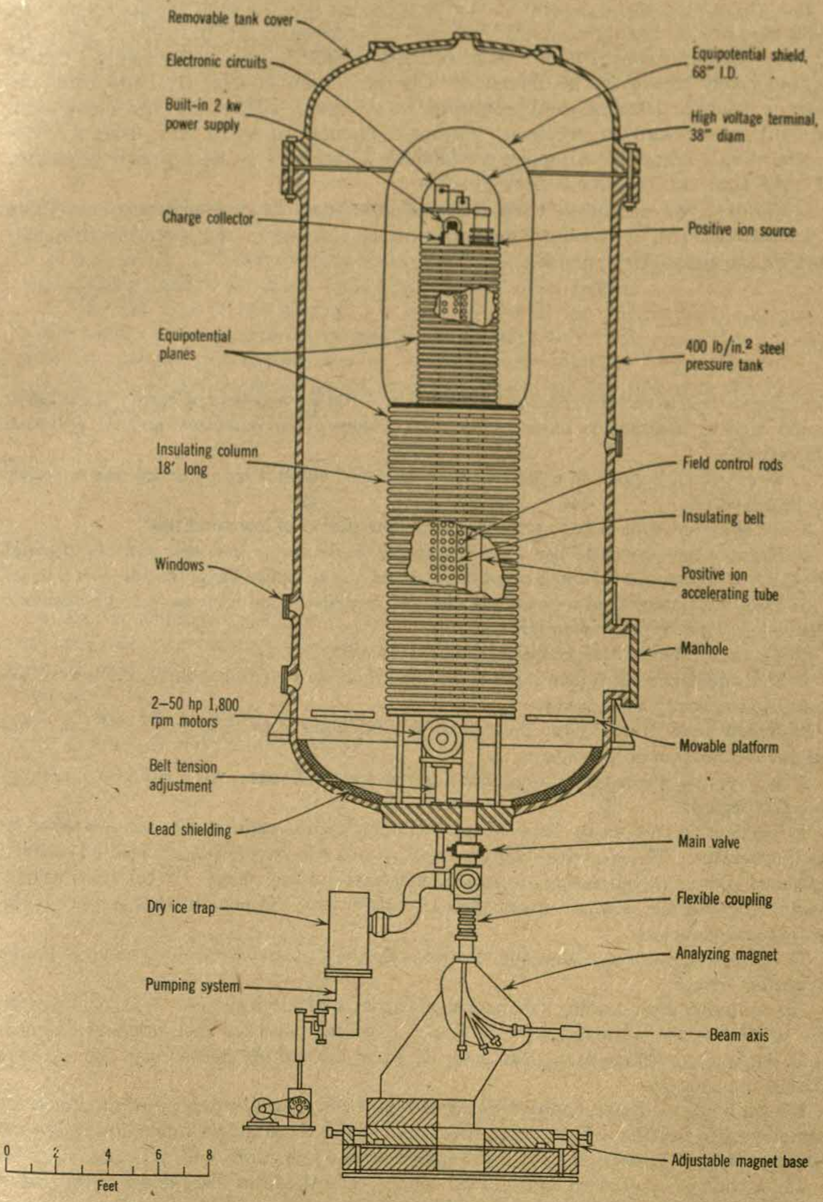


Fig. 29–21 An electrostatic generator at MIT capable of producing 9-Mev protons. The proton beam is accelerated vertically downward, being deflected into a horizontal plane by the analyzing magnet shown at the bottom. (Courtesy of J. G. Trump.)

From another point of view, we note that since the spheres when electrically connected form a single conductor at electrostatic equilibrium there can be only a single potential. This means that $V_r - V_R = 0$, which can occur only if $q = 0$.

In actual electrostatic generators charge is carried into the shell on rapidly moving belts made of insulating material. Charge is "sprayed" onto the belts outside the shell by corona discharge from a series of sharp metallic points connected to a source of moderately high potential difference. Charge is removed from the belts inside the shell by a similar series of points connected to the shell. Electrostatic generators can be built commercially to accelerate protons to energies up to 10 Mev, using a single acceleration. Figure 29–21 shows a schematic diagram of an electrostatic generator at MIT that can produce 9-Mev protons.

Generators can be built in which the accelerated particles are subject to successive accelerations. An electrostatic generator at the University of Pittsburgh subjects particles to three successive accelerations.

QUESTIONS

1. Are we free to call the potential of the earth +100 volts instead of zero? What effect would such an assumption have on measured values of (*a*) potentials and (*b*) potential differences?

2. What would happen to a person on an insulated stand if his potential was increased by 10,000 volts?

3. Do electrons tend to go to regions of high potential or of low potential?

4. Suppose that the earth has a net charge that is *not* zero. Is it still possible to adopt the earth as a standard reference point of potential and to assign the potential $V = 0$ to it?

5. Does the potential of a positively charged insulated conductor have to be positive? Give an example to prove your point.

6. Can two different equipotential surfaces intersect?

7. If **E** equals zero at a given point, must V equal zero for that point? Give some examples to prove your answer.

8. If you know **E** at a given point, can you calculate V at that point? If not, what further information do you need?

9. If V equals a constant throughout a given region of space, what can you say about **E** in that region?

10. In Section 16–6 we saw that the gravitational field strength is zero inside a spherical shell of matter. The electrical field strength is zero not only inside an isolated charged spherical conductor but inside an isolated conductor of *any* shape. Is the gravitational field strength inside, say, a cubical shell of matter zero? If not, in what respect is the analogy not complete?

11. How can you insure that the electric potential in a given region of space will have a constant value?

12. An isolated conducting spherical shell carries a negative charge. What will happen if a positively charged metal object is placed in contact with the shell interior? Assume that the positive charge is (*a*) less than, (*b*) equal to, and (*c*) greater than the negative charge in magnitude.

13. An uncharged metal sphere suspended by a silk thread is placed in a uniform external electric field **E**. What is the magnitude of the electric field for points inside the sphere? Is your answer changed if the sphere carries a charge?

14. A charge is placed on an insulated conductor in the form of a perfect cube. What will be the relative charge density at various points on the cube (surfaces, edges, corners); what will happen to the charge if the cube is in air?

PROBLEMS

1. An infinite charged sheet has a surface charge density σ of 1.0×10^{-7} coul/meter². How far apart are the equipotential surfaces whose potentials differ by 5.0 volts?

2. A charge q is distributed uniformly throughout a nonconducting spherical volume of radius R. (*a*) Show that the potential a distance a from the center, where $a < R$, is given by

$$V = \frac{q(3R^2 - a^2)}{8\pi\epsilon_0 R^3}.$$

(*b*) Is it reasonable that, according to this expression, V is not zero at the center of the sphere?

3. A charge of 10^{-8} coul can be produced by simple rubbing. To what potential would such a charge raise an insulated conducting sphere of 10-cm radius?

4. Consider a point charge with $q = 1.5 \times 10^{-8}$ coul. (*a*) What is the radius of an equipotential surface having a potential of 30 volts? (*b*) Are surfaces whose potentials differ by a constant amount (say 1.0 volt) evenly spaced in radius?

5. In Fig. 29–22, locate the points (*a*) where $V = 0$ and (*b*) where $\mathbf{E} = 0$. Consider only points on the axis and choose $d = 1.0$ meter.

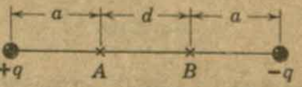


Fig. 29–22

6. In Fig. 29–22 (see Problem 5) sketch qualitatively (*a*) the lines of force and (*b*) the intersections of the equipotential surfaces with the plane of the figure. (Hint: Consider the behavior close to each point charge and at considerable distances from the pair of charges.)

7. In Fig. 29–23 derive an expression for $V_A - V_B$. Does your result reduce to the expected answer when $d = 0$? When $q = 0$?

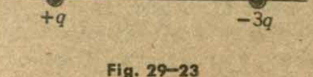


Fig. 29–23

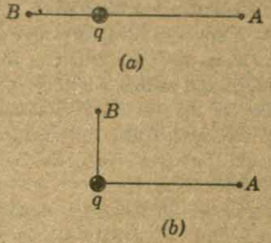


Fig. 29–24

8. A point charge has $q = +1.0 \times 10^{-6}$ coul. Consider point A which is 2.0 meters distant and point B which is 1.0 meter distant in a direction diametrically opposite, as in Fig. 29–24*a*. (*a*) What is the potential difference $V_A - V_B$? (*b*) Repeat if points A and B are located as in Fig. 29–24*b*.

9. Calculate the dipole moment of a water molecule under the assumption that all ten electrons in the molecule circulate symmetrically about the oxygen atom, that the OH distance is 0.96×10^{-8} cm, and that the angle between the two OH bonds is 104°. Compare with the value quoted on p. 719; see Fig. 29–11.

10. For the charge configuration of Fig. 29-25, show that $V(r)$ for points on the vertical axis, assuming $r \gg a$, is given by

$$V = \frac{1}{4\pi\epsilon_0}\left(\frac{q}{r} + \frac{2qa}{r^2}\right).$$

Is this an expected result? (Hint: The charge configuration can be viewed as the sum of an isolated charge and a dipole.)

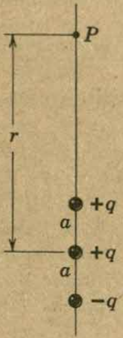


Fig. 29–25

11. In a typical lightning flash the potential difference between discharge points is about 10^9 volts and the quantity of charge transferred is about 30 coul. How much ice would it melt at 0°C if all the energy released could be used for this purpose?

12. Calculate (*a*) the electric potential established by the nucleus of a hydrogen atom at the mean distance of the circulating electron ($r = 5.3 \times 10^{-11}$ meter), (*b*) the electric potential energy of the atom when the electron is at this radius, and (*c*) the kinetic energy of the electron, assuming it to be moving in a circular orbit of this radius centered on the nucleus. (*d*) How much energy is required to ionize the hydrogen atom? Express all energies in electron volts.

13. What is the electric potential energy of the charge configuration of Fig. 29-7? Use the numerical values of Example 5.

14. (*a*) A spherical drop of water carrying a charge of 3×10^{-6} coul has a potential of 500 volts at its surface. What is the radius of the drop? (*b*) If two such drops, of the same charge and radius, combine to form a single spherical drop, what is the potential at the surface of the new drop so formed?

15. If the earth had a net charge equivalent to 1 electron/meter^2 of surface area, (*a*) what would the earth's potential be? (*b*) What would the electric field due to the earth be just outside its surface?

16. Figure 29-26 shows an idealized representation of a U^{238} nucleus ($Z = 92$) on the verge of fission. Calculate (*a*) the repulsive force acting on each fragment and (*b*) the mutual electric potential energy of the two fragments. Assume that the fragments are

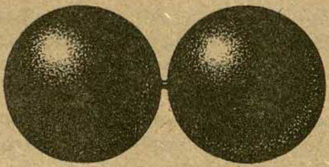


Fig. 29–26

equal in size and charge, spherical, and just touching. The radius of the initially spherical U^{238} nucleus is 8.0×10^{-15} meter. Assume that the material out of which nuclei are made has a constant density.

17. In the Millikan oil drop experiment (see Fig. 27–24) an electric field of 1.92×10^5 nt/coul is maintained at balance across two plates separated by 1.50 cm. Find the potential difference between the plates.

18. (*a*) Show that the electric potential at a point on the axis of a ring of charge of radius a, computed directly from Eq. 29–10, is given by

$$V = \frac{1}{4\pi\epsilon_0}\frac{q}{\sqrt{x^2 + a^2}}.$$

(*b*) From this result derive an expression for E at axial points; compare with the direct calculation of E in Example 5, Chapter 27.

19. In Example 6 the potential at an axial point for a charged disk was shown to be

$$V = \frac{\sigma}{2\epsilon_0}\left(\sqrt{a^2 + r^2} - r\right).$$

From this result show that E for axial points is given by

$$E = \frac{\sigma}{2\epsilon_0}\left(1 - \frac{r}{\sqrt{a^2 + r^2}}\right).$$

Does this expression for E reduce to an expected result for (*a*) $r \gg a$ and (*b*) for $r = 0$?

20. (*a*) Starting from Eq. 29–11, find the magnitude E_r of the radial component of the electric field due to a dipole. (*b*) For what values of θ is E_r zero?

21. Can a conducting sphere 10 cm in radius hold a charge of 4×10^{-6} coul in air without breakdown? The dielectric strength (minimum field required to produce breakdown) of air at 1 atm is 3×10^6 volts/meter.

22. A Geiger counter has a metal cylinder 2.0 cm in diameter along whose axis is stretched a wire 0.005 in. in diameter. If 850 volts are applied between them, what is the electric field strength at the surface of (*a*) the wire and (*b*) the cylinder?

23. Two metal spheres are 3.0 cm in radius and carry charges of $+1.0 \times 10^{-8}$ coul and -3.0×10^{-8} coul, respectively, assumed to be uniformly distributed. If their centers are 2.0 meters apart, calculate (*a*) the potential of the point halfway between their centers and (*b*) the potential of each sphere.

24. In Fig. 29–19 let $R_1 = 1.0$ cm and $R_2 = 2.0$ cm. Before the spheres are connected by the fine wire, a charge of 2.0×10^{-7} coul is placed on the smaller sphere, the larger sphere being uncharged. Calculate (*a*) the charge, (*b*) the charge density, and (*c*) the potential for each sphere after they are connected.

25. The metal object in Fig. 29–27 is a figure of revolution about the horizontal axis. If it is charged negatively, sketch roughly a few equipotentials and lines of force. Use physical reasoning rather than mathematical analysis.

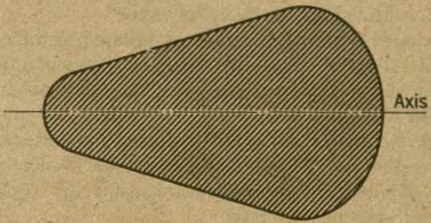


Fig. 29–27

26. Devise an arrangement of three point charges, separated by finite distances, that has zero electric potential energy.

27. Derive an expression for the work required to put the four charges together as indicated in Fig. 29–28.

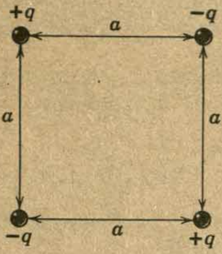


Fig. 29–28

28. A gold nucleus contains a positive charge equal to that of 79 protons. An α-particle ($Z = 2$) has a kinetic energy K at points far from this charge and is traveling directly toward the charge. The particle just touches the surface of the charge (assumed spherical) and is reversed in direction. Calculate K, assuming a nuclear radius of 5.0×10^{-15} meter. The actual α-particle energy used in the experiments of Rutherford and his collaborators was 5.0 Mev. What do you conclude?

29. What is the potential gradient, in volts/meter, at a distance of 10^{-12} meter from the center of the gold nucleus? What is the gradient at the nuclear surface?

30. For the spheres of Fig. 29–19, what is the ratio of electric field strengths at the surface?

31. (*a*) Through what potential difference must an electron fall, according to Newtonian mechanics, to acquire a speed v equal to the speed c of light? (*b*) Newtonian mechanics fails as $v \to c$. Therefore, using the correct relativistic expression for the kinetic energy

$$K = mc^2\left[\frac{1}{\sqrt{1 - (v/c)^2}} - 1\right]$$

in place of the Newtonian expression $K = \frac{1}{2}mv^2$, determine the actual electron speed acquired in falling through the potential difference computed in (*a*). Express this speed as an appropriate fraction of the speed of light.

32. Two insulated concentric conducting spheres of radii R_1 and R_2 carry charges q_1 and q_2, respectively. Derive expressions for $E(r)$ and $V(r)$, where r is the distance from the center of the spheres. Plot $E(r)$ and $V(r)$ from $r = 0$ to $r = 4.0$ meters for $R_1 = 0.50$ meter, $R_2 = 1.0$ meter, $q_1 = +2.0 \times 10^{-6}$ coul, and $q_2 = +1.0 \times 10^{-6}$ coul. Compare with Fig. 29–18.

33. Let the potential difference between the shell of an electrostatic generator and the point at which charges are sprayed onto the moving belt be 3.0×10^6 volts. If the belt transfers charge to the shell at the rate of 3.0×10^{-3} coul/sec, what power must be provided to drive the belt, considering only electrical forces?

34. (*a*) How much charge is required to raise an isolated metallic sphere of 1.0-meter radius to a potential of 1.0×10^6 volts? Repeat for a sphere of 1.0-cm radius. (*b*) Why use a large sphere in an electrostatic generator since the same potential can be achieved for a smaller charge with a small sphere?

35. An alpha particle is accelerated through a potential difference of one million volts in an electrostatic generator. (*a*) What kinetic energy does it acquire? (*b*) What kinetic energy would a proton acquire under these same circumstances? (*c*) Which particle would acquire the greater speed, starting from rest?

Capacitors and Dielectrics

CHAPTER 30

30–1 Capacitance

In Section 29–3 we showed that the potential of a charged conducting sphere, assumed to be completely isolated with no other bodies (conducting or nonconducting) nearby, is given by

$$V_+' = \frac{1}{4\pi\epsilon_0}\frac{q}{R}, \tag{30–1}$$

in which q is the charge on the sphere and R is the sphere radius. The subscript on V indicates that we assume the charge to be positive. We represent this potential in Fig. 30–1 by the line marked V_+'. The line marked V_∞ in that figure represents the potential of an infinitely distant reference position; it has been assigned the value zero, following the usual convention.

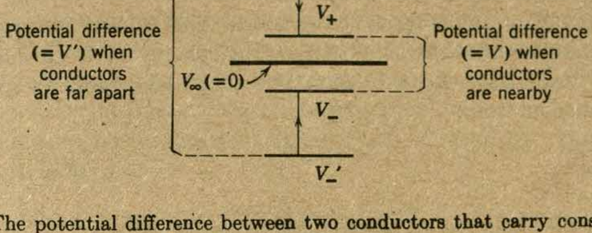

V_+'
V_+
Potential difference ($=V'$) when conductors are far apart
$V_\infty(=0)$
Potential difference ($=V$) when conductors are nearby
V_-
V_-'

Fig. 30–1 The potential difference between two conductors that carry constant, equal, and opposite charges is reduced as the conductors are brought closer together.

Let us now imagine a second sphere of radius R, carrying a negative charge $-q$ and located a large distance ($\gg R$) from the first sphere so that each may still be considered to be electrically isolated. The potential of the second sphere is given by

$$V_-' = -\frac{1}{4\pi\epsilon_0}\frac{q}{R}, \tag{30-2}$$

and this quantity also is represented in Fig. 30–1.

The potential difference V' between the two spheres is

$$V' = V_+' - V_-' = \frac{1}{4\pi\epsilon_0}\frac{2q}{R}.$$

This shows that V', the potential difference, and q, the magnitude of the charge on either sphere, are proportional to each other. We may rewrite the equation as

$$q = (2\pi\epsilon_0 R)V' = C'V', \tag{30-3}$$

in which the proportionality constant in parentheses is called the *capacitance* of the two spheres and assigned the symbol C'.

Let us move the two spheres close together. The presence of each will now spoil the spherical symmetry of the lines of force emanating from the other. Lines from a given sphere, which, for large sphere separations, radiated uniformly in all directions to infinity, now terminate, in part, on the other sphere. Under these conditions Eqs. 30–1 and 30–2 no longer apply, since they were derived (see Section 29–3) on the assumption that spherical symmetry existed, permitting the useful application of Gauss's law.

A positive charge brought near to an isolated object serves to raise the potential of that object and a negative charge serves to lower it, as the student can see by considering the work required to move a positive test charge from infinity to points near such charges. Thus the potential of the positively charged sphere will be lowered by the presence nearby of the negatively charged sphere, from V_+' to some lower value V_+. Similarly the potential of the negative sphere will be raised from V_-' to a higher value V_-. These new potentials are shown in Fig. 30–1, the potential changes for each sphere being suggested by the vertical arrows.

From Fig. 30–1, although the *charges* on the spheres have not changed, it is clear that the *potential difference* between the spheres has been considerably reduced. Put another way, the *capacitance* of the system of two spheres (see Eq. 30–3), defined from

$$C = \frac{q}{V}, \tag{30-4}$$

has been made considerably larger than its initial value C' by bringing the spheres closer together.

It is also possible to use Eq. 30–4 to define the capacitance of a single isolated conductor such as a sphere. In such cases one may imagine that the second "plate," carrying an equal and opposite charge, is a conducting sphere of very large—essentially infinite—radius centered about the conductor. The potential of this infinitely distant sphere, according to the usual convention for potential measurements, is zero. The capacitance of an isolated sphere of radius R is given from Eqs. 30–4 and 30–1 as

$$C = \frac{q}{V} = 4\pi\epsilon_0 R.$$

The mks unit of capacitance that follows from Eq. 30–4 is the coul/volt. A special unit, the *farad*, is used to represent it. It is named in honor of Michael Faraday who, among other contributions, developed the concept of capacitance. Thus

$$1 \text{ farad} = 1 \text{ coul/volt}.$$

The submultiples of the farad, the *microfarad* ($1\ \mu f = 10^{-6}$ farad) and the *micromicrofarad* ($1\ \mu\mu f = 10^{-12}$ farad), are more convenient units in practice.

An analogy can be made between a capacitor carrying a charge q and a rigid container of volume $\mathcal{V}$ containing μ moles of an ideal gas.

The gas pressure p is directly proportional to μ for a fixed temperature, according to the ideal gas law (Eq. 23–2)

$$\mu = \left(\frac{\mathcal{V}}{RT}\right) p.$$

For the capacitor (Eq. 30–4)

$$q = (C)V.$$

Comparison shows that the capacitance C of the capacitor, assuming a fixed temperature, is analogous to the volume $\mathcal{V}$ of the container.

Note that any amount of charge can be put on the capacitor, and any mass of gas can be put in the container, up to certain limits. These correspond to electrical breakdown ("arcing over") for the capacitor and to rupture of the walls for the container.

Figure 30–2 shows a more general case of two nearby conductors, which are now permitted to be of any shape, carrying equal and opposite charges. Such an arrangement is called a *capacitor*, the conductors being called *plates*. The equal and opposite charges might be established by connecting the plates momentarily to opposite poles of a battery. The capacitance C of any capacitor is defined from Eq. 30–4 in which we remind the student that V is the *potential difference between the plates* and q is the magnitude of the *charge on either plate*; q must not be taken as the net charge of the capacitor, which is zero. The capacitance of a capacitor depends on the geometry of each plate, their spatial relationship to each other, and the medium in which the plates are immersed. For the present, we take this medium to be a vacuum.

Capacitors are very useful devices, of great interest to physicists and engineers. For example:

1. In this book we stress the importance of *fields* to the understanding of natural phenomena. A capacitor can be used to establish desired electric field configurations for various purposes. In Section 27–5 we described the deflection of an electron beam in a uniform field set up by a capacitor, al-

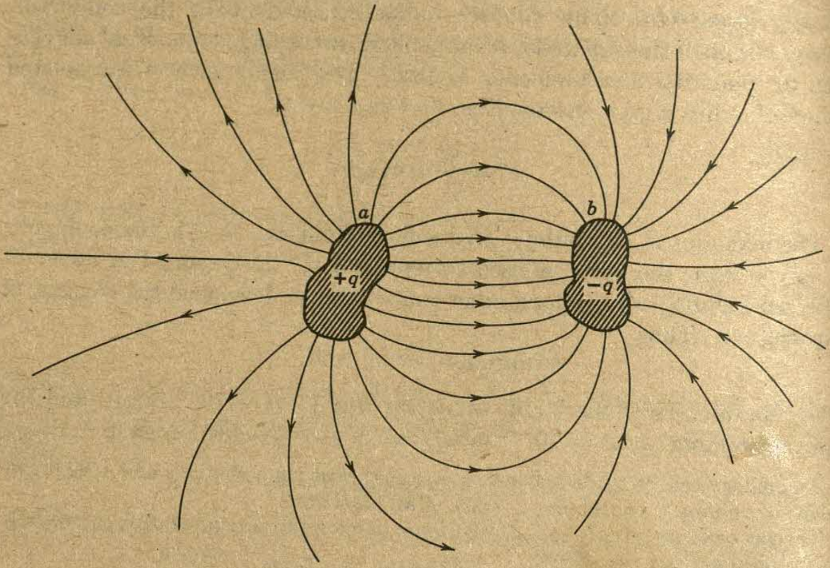


Fig. 30–2 Two insulated conductors carrying equal and opposite charges form a capacitor.

though we did not use this term in that section. In later sections we discuss the behavior of dielectric materials when placed in an electric field (provided conveniently by a capacitor) and we shall see how the laws of electromagnetism can be generalized to take the presence of dielectric bodies more readily into account.

2. A second important concept stressed in this book is *energy*. By analyzing a charged capacitor we show that electric energy may be considered to be stored in the electric field between the plates and indeed in any electric field, however generated. Because capacitors can confine strong electric fields to small volumes, they can serve as useful devices for storing energy. In many electron synchrotrons, which are cyclotron-like devices for accelerating electrons, energy accumulated and stored in a large bank of capacitors over a relatively long period of time is made available intermittently to accelerate the electrons by discharging the capacitor in a much shorter time. Many researches and devices in plasma physics also make use of bursts of energy stored in this way.

3. The electronic age could not exist without capacitors. They are used, in conjunction with other devices, to reduce voltage fluctuations in electronic power supplies, to transmit pulsed signals, to generate or detect electromagnetic oscillations at radio frequencies, and to provide time delays. In most of these applications the potential difference between the plates will not be constant, as we assume in this chapter, but will vary with time, often in a sinusoidal or a pulsed fashion. In later chapters we consider some aspects of the capacitor used as a circuit element.

30–2 Calculating Capacitance

Figure 30–3 shows a *parallel-plate* capacitor formed of two parallel conducting plates of area A separated by a distance d. If we connect each plate to the terminal of a battery, a charge $+q$ will appear on one plate and a charge $-q$ on the other. If d is small compared with the plate dimensions, the electric field strength **E** between the plates will be uniform, which means that the lines of force will be parallel and evenly spaced. The laws of electromagnetism (see Problem 20, Chapter 35) require that there be some "fringing" of the lines at the edges of the plates, but for small enough d it can be neglected for our present purpose.

We can calculate the capacitance of this device using Gauss's law. Figure 30–3 shows (dashed lines) a Gaussian surface of height h closed by plane caps of area A that are the shape and size of the capacitor plates. The flux of **E** is zero for the part of the Gaussian surface that lies inside the top capacitor plate because the electric field inside a conductor carrying a static charge is zero. The flux of **E** through the wall of the Gaussian surface is zero because, to the extent that the fringing of the lines of force can be neglected, **E** lies in the wall.

This leaves only the face of the Gaussian surface that lies between the plates. Here **E** is constant and the flux Φ_E is simply EA. Gauss's law gives

$$\epsilon_0\Phi_E = \epsilon_0 EA = q. \tag{30–5}$$

The work required to carry a test charge q_0 from one plate to the other can be expressed either as q_0V (see Eq. 29–1) or as the product of a force q_0E times a distance d or q_0Ed. These expressions must be equal, or

$$V = Ed. \tag{30–6}$$

More formally, Eq. 30–6 is a special case of the general relation (Eq. 29–5; see also Example 1, Chapter 29)

$$V = -\int \mathbf{E}\cdot d\mathbf{l},$$

where V is the difference in potential between the plates. The integral may

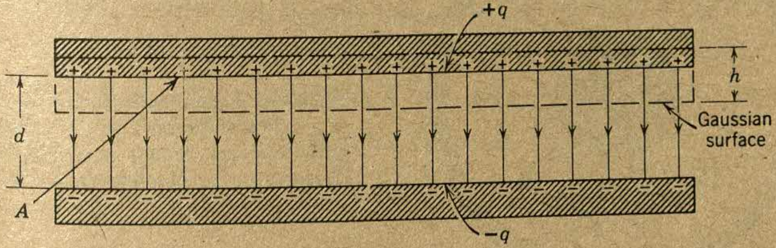


Fig. 30–3 A parallel-plate capacitor with plates of area A. The dashed line represents a Gaussian surface whose height is h and whose top and bottom caps are the same shape and size as the capacitor plates.

be taken over any path that starts on one plate and ends on the other because each plate is an equipotential surface and the electrostatic force is path independent. Although the simplest path between the plates is a perpendicular straight line, Eq. 30–6 follows no matter what path of integration we choose.

If we substitute Eqs. 30–5 and 30–6 into the relation $C = q/V$, we obtain

$$C = \frac{q}{V} = \frac{\epsilon_0 EA}{Ed} = \frac{\epsilon_0 A}{d}. \tag{30–7}$$

Equation 30–7 holds only for capacitors of the parallel-plate type; different formulas hold for capacitors of different geometry.

In Section 26–4 we stated that ϵ_0, which we first met in connection with Coulomb's law, was not measured in terms of that law because of experimental difficulties. Equation 30–7 suggests that ϵ_0 might be measured by building a capacitor of accurately known plate area and plate spacing and determining its capacitance experimentally by measuring q and V in the relation $C = q/V$. Thus Eq. 30–7 can be solved for ϵ_0 and a numerical value found in terms of the measured quantities A, d, and C; ϵ_0 has been measured accurately in this way.

▶ **Example 1.** The parallel plates of an air-filled capacitor are everywhere 1.0 mm apart. What must the plate area be if the capacitance is to be 1.0 farad?

From Eq. 30–7

$$A = \frac{dC}{\epsilon_0} = \frac{(1.0 \times 10^{-3}\ \text{meter})(1.0\ \text{farad})}{8.9 \times 10^{-12}\ \text{coul}^2/\text{nt-m}^2} = 1.1 \times 10^8\ \text{meter}^2.$$

This is the area of a square sheet more than 6 miles on edge; the farad is indeed a large unit.

Example 2. *A cylindrical capacitor.* A cylindrical capacitor consists of two coaxial cylinders (Fig. 30–4) of radius a and b and length l. What is the capacitance of this device? Assume that the capacitor is very long (that is, that $l \gg b$) so that fringing of the lines of force at the ends can be ignored for the purpose of calculating the capacitance.

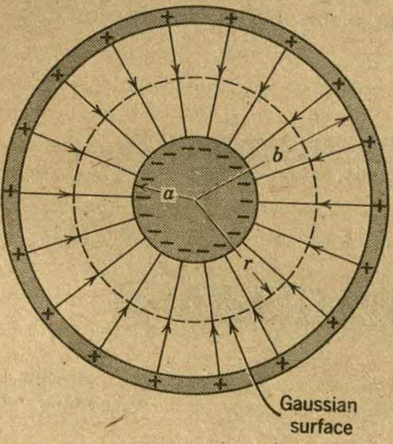


Fig. 30–4 Example 2. A cross section of a cylindrical capacitor. The dashed circle is a cross section of a cylindrical Gaussian surface of radius r and length l.

As a Gaussian surface construct a coaxial cylinder of radius r and length l, closed by plane caps. Gauss's law

$$\epsilon_0 \oint \mathbf{E} \cdot d\mathbf{S} = q$$

gives

$$\epsilon_0 E(2\pi r)(l) = q,$$

the flux being entirely through the cylindrical surface and not through the end caps. Solving for E yields

$$E = \frac{q}{2\pi\epsilon_0 rl}.$$

The potential difference between the plates is given by Eq. 29–5 [note that $\mathbf{E}$ an $d\mathbf{l}$ ($= d\mathbf{r}$) point in opposite directions] or

$$V = -\int_a^b \mathbf{E} \cdot d\mathbf{l} = \int_a^b E\,dr = \int_a^b \frac{q}{2\pi\epsilon_0 l}\frac{dr}{r} = \frac{q}{2\pi\epsilon_0 l}\ln\frac{b}{a}.$$

Finally, the capacitance is given by

$$C = \frac{q}{V} = \frac{2\pi\epsilon_0 l}{\ln(b/a)}.$$

Like the relation for the parallel-plate capacitor (Eq. 30–7), this relation also depend only on geometrical factors.

Example 3. *Capacitors in parallel.* Figure 30–5 shows three capacitors connected in parallel. What single capacitance C is equivalent to this combination? "Equivalent" means that if the parallel combination and the single capacitance were each in a box with wires a and b connected to terminals, it would not be possible to distinguish the two by electrical measurements external to the box.

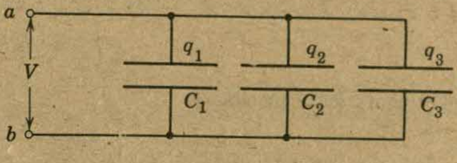


Fig. 30–5 Example 3. Three capacitors in parallel.

The potential difference across each capacitor in Fig. 30–5 will be the same. This follows because all of the upper plates are connected together and to terminal a, whereas all of the lower plates are connected together and to terminal b. Applying the relation $q = CV$ to each capacitor yields

$$q_1 = C_1 V; \qquad q_2 = C_2 V; \qquad \text{and} \qquad q_3 = C_3 V.$$

The total charge q on the combination is

$$\begin{aligned} q &= q_1 + q_2 + q_3 \\ &= (C_1 + C_2 + C_3)V. \end{aligned}$$

The equivalent capacitance C is

$$C = \frac{q}{V} = C_1 + C_2 + C_3.$$

result can easily be extended to any number of parallel-connected capacitors.

Example 4. *Capacitors in series.* Figure 30–6 shows three capacitors connected in series. What single capacitance C is "equivalent" (see Example 3) to this combination?

For capacitors connected as shown, the magnitude q of the charge on each plate must be the same. This is true because the net charge on the part of the circuit enclosed by the dashed line in Fig. 30–6 must be zero; that is, the charge present on these plates initially is zero and connecting a battery between a and b will only produce a charge separation, the *net* charge on these plates still being zero. Assuming that neither C_1 nor C_2 "sparks over," there is no way for charge to enter or leave the region enclosed by the dashed line.

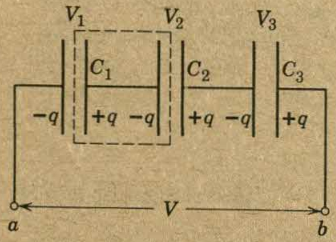


Fig. 30–6 Example 4. Three capacitors in series.

Applying the relation $q = CV$ to each capacitor yields

$$V_1 = q/C_1; \qquad V_2 = q/C_2; \qquad \text{and} \qquad V_3 = q/C_3.$$

The potential difference for the series combination is

$$\begin{aligned} V &= V_1 + V_2 + V_3 \\ &= q\left(\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3}\right). \end{aligned}$$

The equivalent capacitance

$$C = \frac{q}{V} = \frac{1}{\dfrac{1}{C_1} + \dfrac{1}{C_2} + \dfrac{1}{C_3}},$$

or

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3}.$$

The equivalent series capacitance is always less than the smallest capacitance in the chain. ◀

30–3 Parallel-Plate Capacitor with Dielectric

Equation 30–7 holds only for a parallel-plate capacitor with its plates in a vacuum. Michael Faraday, in 1837, first investigated the effect of filling the space between the plates with a dielectric, say mica or oil. In Faraday's words:

> The question may be stated thus: suppose A an electrified plate of metal suspended in air, and B and C two exactly similar plates, placed parallel to and on each side of A at equal distances and insulated; A will then induce equally toward B and C [that is, equal charges will appear on these plates]. If in this position of the plates some other dielectric than air, as shell-lac, be introduced between A and

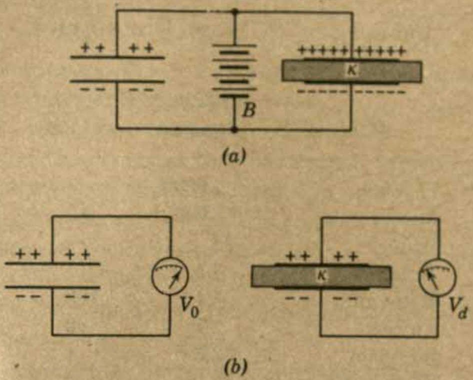


Fig. 30–7 (*a*) Battery *B* supplies the same potential difference to each capacitor; the one on the right has the higher charge. (*b*) Both capacitors carry the same charge; the one on the right has the lower potential difference, as indicated by the meter readings.

> *C*, will the induction between them remain the same? Will the relation of *C* and *B* to *A* be unaltered, notwithstanding the difference of the dielectrics interposed between them?

Faraday answered this question by constructing two identical capacitors, in one of which he placed a dielectric, the other containing air at normal pressure. When both capacitors were charged to the *same potential difference*, Faraday found by experiment that *the charge on the one containing the dielectric was greater than that on the other;* see Fig. 30–7*a*.

Faraday measured the relative charges on the plates of the two capacitors by touching a metal ball (fitted with an insulating handle) to the plates, thus sampling the charge quantitatively. He then put this ball in a Coulomb torsion balance and measured the force of Coulomb repulsion on a second (standard) charged ball mounted on the balance arm.

Since q is larger, for the same V, if a dielectric is present, it follows from the relation $C = q/V$ that *the capacitance of a capacitor increases if a dielectric is placed between the plates.* The ratio of the capacitance with the dielectric * to that without is called the *dielectric constant* κ of the material; see Table 30–1.

Instead of maintaining the two capacitors at the same potential difference, we can place the *same charge* on them, as in Fig. 30–7*b*. Experiment then shows that the potential difference V_d between the plates of the right-hand capacitor is smaller than that for the left-hand capacitor by the factor $1/\kappa$, or

$$V_d = V_0/\kappa.$$

We are led once again to conclude, from the relation $C = q/V$, that the effect of the dielectric is to increase the capacitance by a factor κ.

---

* Assumed to fill completely the space between the plates.

Table 30–1

PROPERTIES OF SOME DIELECTRICS

Material	Dielectric Constant	Dielectric Strength * (kv/mm)
Vacuum	1.00000	∞
Air	1.00054	0.8
Water	78	–
Paper	3.5	14
Ruby mica	5.4	160
Amber	2.7	90
Porcelain	6.5	4
Fused quartz	3.8	8
Pyrex glass	4.5	13
Bakelite	4.8	12
Polyethylene	2.3	50
Polystyrene	2.6	25
Teflon	2.1	60
Neoprene	6.9	12
Pyranol oil	4.5	12
Titanium dioxide	100	6

* This is the maximum potential gradient that may exist in the dielectric without the occurrence of electrical breakdown. Dielectrics are often placed between conducting plates to permit a higher potential difference to be applied between them than would be possible with air as the dielectric.

For a parallel-plate capacitor we can write, as an experimental result,

$$C = \frac{\kappa\epsilon_0 A}{d}. \tag{30–8}$$

Equation 30–7 is a special case of this relation found by putting $\kappa = 1$, corresponding to a vacuum between the plates. Experiment shows that the capacitance of *all* types of capacitor is increased by the factor κ if the space between the plates is filled with a dielectric. Thus the capacitance of any capacitor can be written as

$$C = \kappa\epsilon_0 L,$$

where L depends on the geometry and has the dimensions of a length. For a parallel-plate capacitor (see Eq. 30–7) L is A/d; for a cylindrical capacitor (see Example 2) it is $2\pi l/\ln(b/a)$.

30–4 Dielectrics—An Atomic View

We now seek to understand, in atomic terms, what happens when a dielectric is placed in an electric field. There are two possibilities. The molecules of some dielectrics, like water, have permanent electric dipole moments.

In such materials (called *polar*) the electric dipole moments **p** tend to align themselves with an external electric field, as in Fig. 30-8; see also Section 27-6. Because the molecules are in constant thermal agitation, the degree of alignment will not be complete but will increase as the applied electric field is increased or as the temperature is decreased.

Whether or not the molecules have permanent electric dipole moments, they acquire them by *induction* when placed in an electric field. In Section 29-5 we saw that the external electric field tends to separate the negative and the positive charge in the atom or molecule. This *induced electric dipole moment* is present only when the electric field is present. It is proportional to the electric field (for normal field strengths) and is created already lined up with the electric field as Fig. 29-12 suggests.

Let us use a parallel-plate capacitor, carrying a fixed charge q and not connected to a battery (see Fig. 30-7*b*), to provide a uniform external electric field $\mathbf{E}_0$ into which we place a dielectric slab. The over-all effect of alignment and induction is to separate the center of positive charge of the entire slab slightly from the center of negative charge. The slab, as a whole, although remaining electrically neutral, becomes *polarized*, as Fig. 30-9*b* suggests. The net effect is a pile-up of positive charge on the right face of the slab and of negative charge on the left face; within the slab no excess charge appears in any given volume element. Since the slab as a whole remains neutral, the positive *induced surface charge* must be equal in magnitude to the negative induced surface charge. Note that in this process electrons in the dielectric

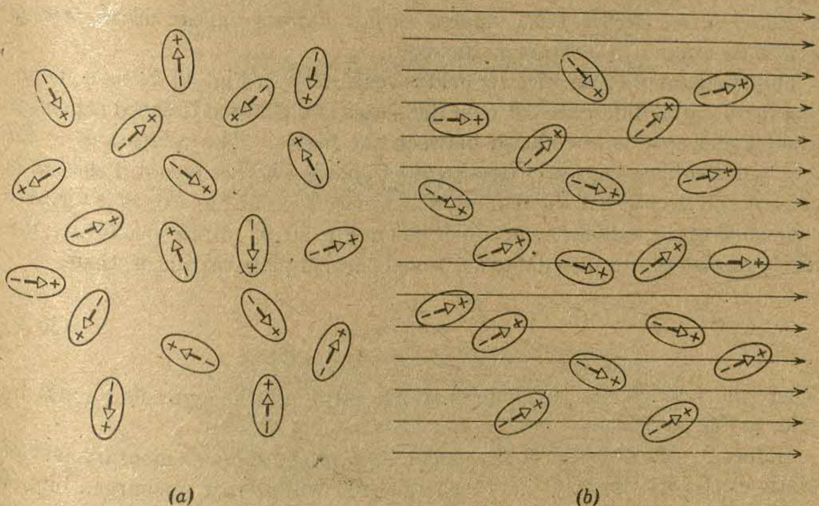


Fig. 30-8 (*a*) Molecules with a permanent electric dipole moment, showing their random orientation in the absence of an external electric field. (*b*) An electric field is applied, producing partial alignment of the dipoles. Thermal agitation prevents complete alignment.

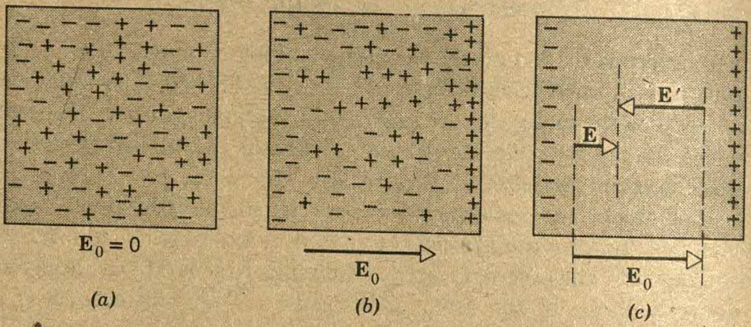


Fig. 30–9 (*a*) A dielectric slab, showing the random distribution of plus and minus charges. (*b*) An external field $\mathbf{E}_0$, established by putting the slab between the plates of a parallel-plate capacitor (not shown), separates the center of plus charge in the slab slightly from the center of minus charge, resulting in the appearance of surface charges. No *net* charge exists in any volume element located in the *interior* of the slab. (*c*) The surface charges set up a field $\mathbf{E}'$ which opposes the external field $\mathbf{E}_0$ associated with the charges on the capacitor plates. The resultant field $\mathbf{E}$ ($= \mathbf{E}_0 + \mathbf{E}'$) in the dielectric is thus less than $\mathbf{E}_0$.

are displaced from their equilibrium positions by distances that are considerably less than an atomic diameter. There is no transfer of charge over macroscopic distances such as occurs when a current is set up in a conductor.

Figure 30–9*c* shows that the induced surface charges will always appear in such a way that the electric field set up *by them* ($\mathbf{E}'$) opposes the external electric field $\mathbf{E}_0$. The *resultant* field in the dielectric $\mathbf{E}$ is the vector sum of $\mathbf{E}_0$ and $\mathbf{E}'$. It points in the same direction as $\mathbf{E}_0$ but is smaller. *If a dielectric is placed in an electric field, induced surface charges appear which tend to weaken the original field within the dielectric.*

This weakening of the electric field reveals itself in Fig. 30–7*b* as a reduction in potential difference between the plates of a charged isolated capacitor when a dielectric is introduced between the plates. The relation $V = Ed$ for a parallel-plate capacitor (see Eq. 30–6) holds whether or not dielectric is present and shows that the reduction in V described in Fig. 30–7*b* is directly connected to the reduction in E described in Fig. 30–9. More specifically, if a dielectric slab is introduced into a charged parallel-plate capacitor, then

$$\frac{E_0}{E} = \frac{V_0}{V_d} = \kappa \tag{30–9}$$

where the symbols on the left refer to Fig. 30–9 and the symbols V_0 and V_d refer to Fig. 30–7*b*.*

Induced surface charge is the explanation of the most elementary fact of static electricity, namely, that a charged rod will attract uncharged bits of

---

* Equation 30–9 does not hold if the battery remains connected while the dielectric slab is introduced. In this case V (hence E) could not change. Instead, the charge q on the capacitor plates would increase by a factor κ, as Fig. 30–7*a* suggests.

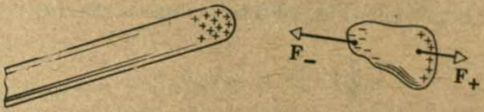


Fig. 30–10 A charged rod attracts an uncharged piece of paper because unbalanced forces act on the induced surface charges.

paper, etc. Figure 30–10 shows a bit of paper in the field of a charged rod. Surface charges appear on the paper as shown. The negatively charged end of the paper will be pulled toward the rod and the positively charged end will be repelled. These two forces do not have the same magnitude because the negative end, being closer to the rod, is in a stronger field and experiences a stronger force. The net effect is an attraction. A dielectric body in a *uniform* electric field will not experience a net force.

30–5 Dielectrics and Gauss's Law

So far our use of Gauss's law has been confined to situations in which no dielectric was present. Now let us apply this law to a parallel-plate capacitor filled with a dielectric of dielectric constant κ.

Figure 30–11 shows the capacitor both with and without the dielectric. It is assumed that the charge q on the plates is the same in each case. Gaussian surfaces have been drawn after the fashion of Fig. 30–3.

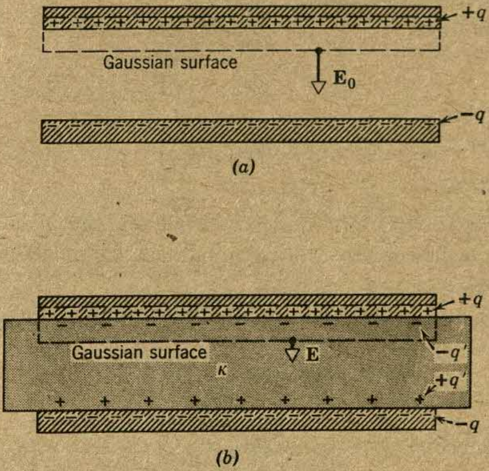


Fig. 30–11 A parallel-plate capacitor (a) without and (b) with a dielectric. The charge q on the plates is assumed to be the same in each case.

If no dielectric is present (Fig. 30–11*a*), Gauss's law (see Eq. 30–5) gives

$$\epsilon_0 \oint \mathbf{E} \cdot d\mathbf{S} = \epsilon_0 E_0 A = q$$

or

$$E_0 = \frac{q}{\epsilon_0 A}. \tag{30–10}$$

If the dielectric is present (Fig. 30–11*b*), Gauss's law gives

$$\epsilon_0 \oint \mathbf{E} \cdot d\mathbf{S} = \epsilon_0 E A = q - q'$$

or

$$E = \frac{q}{\epsilon_0 A} - \frac{q'}{\epsilon_0 A}, \tag{30–11}$$

in which $-q'$, the *induced surface charge*, must be distinguished from q, the so-called *free charge* on the plates. These two charges, both of which lie within the Gaussian surface, are opposite in sign; $q - q'$ is the *net* charge within the Gaussian surface.

Equation 30–9 shows that in Fig. 30–11

$$E = \frac{E_0}{\kappa}.$$

Combining this with Eq. 39–10, we have

$$E = \frac{E_0}{\kappa} = \frac{q}{\kappa \epsilon_0 A}. \tag{30–12}$$

Inserting this in Eq. 30–11 yields

$$\frac{q}{\kappa \epsilon_0 A} = \frac{q}{\epsilon_0 A} - \frac{q'}{\epsilon_0 A} \tag{30–13a}$$

or

$$q' = q\left(1 - \frac{1}{\kappa}\right). \tag{30–13b}$$

This shows correctly that the induced surface charge q' is always less in magnitude than the free charge q and is equal to zero if no dielectric is present, that is, if $\kappa = 1$.

Now we write Gauss's law for the case of Fig. 30–11*b* in the form

$$\epsilon_0 \oint \mathbf{E} \cdot d\mathbf{S} = q - q', \tag{30–14}$$

$q - q'$ again being the net charge within the Gaussian surface. Substituting from Eq. 30–13*b* for q' leads, after some rearrangement, to

$$\epsilon_0 \oint \kappa \mathbf{E} \cdot d\mathbf{S} = q. \tag{30–15}$$

This important relation, although derived for a parallel-plate capacitor, i true generally and is the form in which Gauss's law is usually written when dielectrics are present. Note the following:

1. The flux integral now contains a factor κ.
2. The charge q contained within the Gaussian surface is taken to be the *free charge only*. Induced surface charge is deliberately ignored on the right side of this equation, having been taken into account by the introduction of κ on the left side. Equations 30–14 and 30–15 are completely equivalent formulations.

▶ **Example 5.** Figure 30–12 shows a dielectric slab of thickness b and dielectric constant κ placed between the plates of a parallel-plate capacitor of plate area A and separation d. A potential difference V_0 is applied with no dielectric present. The battery is then disconnected and the dielectric slab inserted. Assume that $A = 100$ cm^2, $d = 1.0$ cm, $b = 0.50$ cm, $\kappa = 7.0$, and $V_0 = 100$ volts and (a) calculate the capacitance C_0 before the slab is inserted.

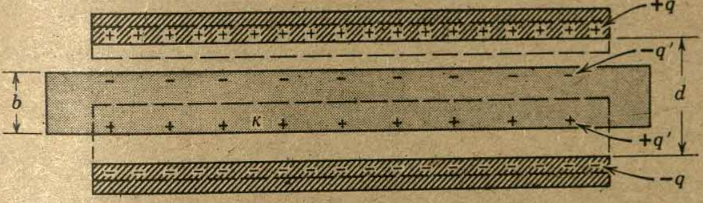


Fig. 30–12 Example 5. A parallel-plate capacitor containing a dielectric slab.

From Eq. 30–7, C_0 is found:

$$C_0 = \frac{\epsilon_0 A}{d} = \frac{(8.9 \times 10^{-12}\ \text{coul}^2/\text{nt-m}^2)(10^{-2}\ \text{meter}^2)}{10^{-2}\ \text{meter.}} = 8.9\ \mu\mu\text{f}.$$

(b) Calculate the free charge q.
From Eq. 30–4,

$$q = C_0 V_0 = (8.9 \times 10^{-12}\ \text{farad})(100\ \text{volts}) = 8.9 \times 10^{-10}\ \text{coul}.$$

Because of the technique used to charge the capacitor, the free charge remains unchanged as the slab is introduced. If the charging battery had *not* been disconnected this would not be the case.

(c) Calculate the electric field strength in the gap.

Applying Gauss's law in the form given in Eq. 30–15 to the Gaussian surface Fig. 30–12 (upper plate) yields

$$\epsilon_0 \oint \kappa \mathbf{E} \cdot d\mathbf{S} = \epsilon_0 E_0 A = q,$$

$$\text{or}\quad E_0 = \frac{q}{\epsilon_0 A} = \frac{8.9 \times 10^{-10}\ \text{coul}}{(8.9 \times 10^{-12}\ \text{coul}^2/\text{nt-m}^2)(10^{-2}\ \text{meter}^2)} = 1.0 \times 10^4\ \text{volts/mete}$$

Note that we put $\kappa = 1$ here because the surface over which we evaluate flux integr does not pass through any dielectric. Note too that E_0 remains unchanged when the slab is introduced; this derivation takes no specific account of the presence of the dielectric.

(*d*) Calculate the electric field strength in the dielectric.

Applying Eq. 30–15 to the Gaussian surface of Fig. 30–12 (lower plate) yields

$$\epsilon_0 \oint \kappa \mathbf{E} \cdot d\mathbf{S} = \epsilon_0 \kappa E A = q.$$

Note that κ appears here because the surface cuts through the dielectric and that only the free charge q appears on the right. Thus we have

$$E = \frac{q}{\kappa \epsilon_0 A} = \frac{E_0}{\kappa} = \frac{1.0 \times 10^4 \text{ volts/meter}}{7.0} = 0.14 \times 10^4 \text{ volts/meter}.$$

(*e*) Calculate the potential difference between the plates.

Applying Eq. 29–5 to a straight perpendicular path from the lower plate (L) to the upper one (U) yields

$$V = -\int_L^U \mathbf{E} \cdot d\mathbf{l} = -\int_L^U E \cos 180^\circ \, dl = \int_L^U E \, dl = E_0(d - b) + Eb.$$

Numerically

$$V = (1.0 \times 10^4 \text{ volts/meter})(5 \times 10^{-3} \text{ meter}) + (0.14 \times 10^4 \text{ volts/meter})(5 \times 10^{-3} \text{ meter}) = 57 \text{ volts}.$$

This contrasts with the original applied potential difference of 100 volts; compare Fig. 30–7*b*.

(*f*) Calculate the capacitance with the slab in place.

From Eq. 30–4,

$$C = \frac{q}{V} = \frac{8.9 \times 10^{-10} \text{ coul}}{57 \text{ volts}} = 16 \,\mu\mu\text{f}.$$

When the dielectric slab is introduced, the potential difference drops from 100 to 57 volts and the capacitance rises from 8.9 to 16 $\mu\mu$f, a factor of 1.8. If the dielectric slab had filled the capacitor, the capacitance would have risen by a factor of κ (= 7.0) to 62 $\mu\mu$f. ◀

30–6 Three Electric Vectors

For all situations that we encounter in this book our discussion of the behavior of dielectrics in an electric field is adequate. However, the problems that we treat are simple ones, such as that of a rectangular slab placed at right angles to a uniform external electric field. For more difficult problems, such as that of finding **E** at the center of a dielectric ellipsoid placed in a (possibly nonuniform) external electric field, it greatly simplifies the labor and leads to deeper insight if we introduce a new formalism. We do so largely so that students who take a second course in electromagnetism will have some familiarity with the concepts.

Let us rewrite Eq. 30–13*a*, which applies to a parallel-plate capacitor containing a dielectric, as

$$\frac{q}{A} = \epsilon_0 \left(\frac{q}{\kappa \epsilon_0 A} \right) + \frac{q'}{A}. \qquad (30\text{–}16)$$

The quantity in parentheses (see Eq. 30–12) is simply the electric field strength E in the dielectric. The last term in Eq. 30–16 is the *induced surface charge per unit area*. We call it the *electric polarization* P, or

$$P = \frac{q}{A}. \qquad (30\text{–}17)$$

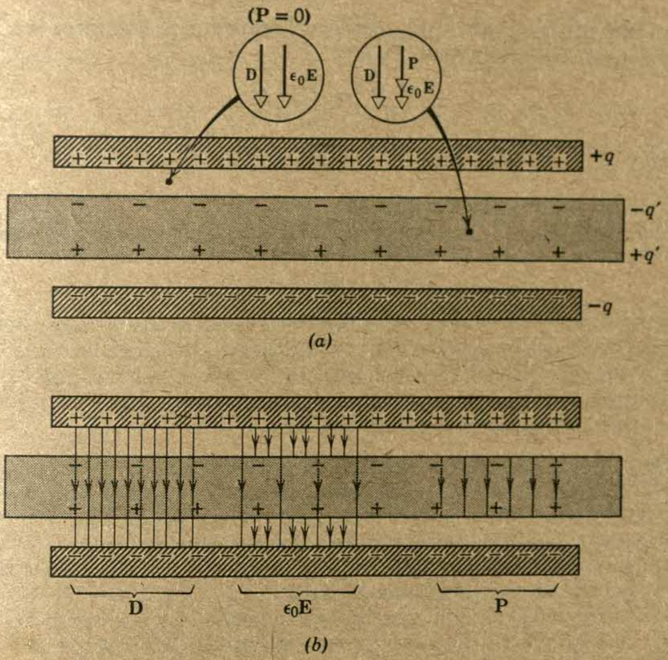


Fig. 30–13 (*a*) Showing **D**, ϵ_0**E**, and **P** in the dielectric (*upper right*) and in the gap (*upper left*) for a parallel-plate capacitor. (*b*) Showing samples of the lines associated with **D** (free charge), ϵ_0**E** (all charges), and **P** (polarization charge).

The name is suitable because the induced surface charge q' appears when the dielectric is polarized.

The electric polarization P can be defined in an equivalent way by multiplying the numerator and denominator in Eq. 30–17 by d, the thickness of the dielectric slab in Fig. 30–11,

$$P = \frac{q'd}{Ad}. \tag{30–18}$$

The numerator is the product $q'd$ of the magnitude of the (equal and opposite) polarization charges by their separation. It is thus the induced electric dipole moment of the dielectric slab. Since the denominator Ad is the volume of the slab, we see that the electric polarization can also be defined as the induced electric dipole moment per unit volume in the dielectric. This definition suggests that since the electric dipole moment is a vector the electric polarization is also a vector, its magnitude being P. The direction of **P** is from the negative induced charge to the positive induced charge, as for any dipole. In Fig. 30–13, which shows a capacitor with a dielectric slab filling half the space between the plates, **P** points down.

We can now rewrite Eq. 30–16 as

$$\frac{q}{A} = \epsilon_0 E + P. \tag{30–19}$$

The quantity on the right occurs so often in electrostatic problems that we give it the special name *electric displacement* D, or

$$D = \epsilon_0 E + P \tag{30-20a}$$

in which

$$D = \frac{q}{A}. \tag{30-20b}$$

The name has historical significance only.

Since **E** and **P** are vectors, **D** must also be one, so that in the more general case we have

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}. \tag{30-21}$$

In Fig. 30–13 all three vectors point down and each has a constant magnitude for every point in the dielectric (and also at every point in the air gap) so that the vector nature of Eq. 30–21 is not very important in this case. In more complicated problems, however, **E**, **P**, and **D** may vary in magnitude and direction from point to point.

From their definitions we see the following:

1. **D** (see Eq. 30–20*b*) is connected with the *free charge* only. We can represent the vector field of **D** by *lines of* **D**, just as we represent the field of **E** by lines of force. Figure 30–13*b* shows that the lines of **D** begin and end on the free charges.
2. **P** (see Eq. 30–17) is connected with the *polarization charge* only. It is also possible to represent this vector field by lines. Figure 30–13*b* shows that the lines of **P** begin and end on the polarization charges.
3. **E** is connected with all charges that are actually present, whether free or polarization. The lines of **E** reflect the presence of both kinds of charge, as Fig. 30–13*b* shows. Note (Eqs. 30–17 and 30–20*b*) that the units for **P** and **D** (coul/meter^2) differ from those of **E** (nt/coul).

The electric field vector **E**, which is what determines the force that acts on a suitably placed test charge, remains of fundamental interest. **D** and **P** are auxiliary vectors useful as aids in the solution of problems more complex than that of Fig. 30–13.

The vectors **D** and **P** can both be expressed in terms of **E** alone. A convenient starting point is the identity

$$\frac{q}{A} = \kappa\epsilon_0 \left(\frac{q}{\kappa\epsilon_0 A}\right).$$

Comparison with Eqs. 30–12 and 30–20*b* shows that this, extended to vector form, can be written as

$$\mathbf{D} = \kappa\epsilon_0 \mathbf{E}. \tag{30-22}$$

We can also write the polarization (see Eqs. 30–17 and 30–13*b*) as

$$P = \frac{q'}{A} = \frac{q}{A}\left(1 - \frac{1}{\kappa}\right).$$

Since $q/A = D$, we can rewrite this, using Eq. 30–22 and casting the result into vector form, as

$$\mathbf{P} = \epsilon_0(\kappa - 1)\mathbf{E}. \tag{30-23}$$

This shows clearly that in a vacuum ($\kappa = 1$) the polarization vector **P** is zero.*

---

* Certain waxes, when polarized in their molten state, retain a permanent polarization after solidifying, even though the external polarizing field is removed. *Electrets*, manufactured in this way, are the electrostatic analog of permanent magnets in that they possess a gross permanent electric dipole moment. Materials from which electrets can be constructed are called *ferroelectric*. Electrets do not obey Eq. 30–23 because they have a nonvanishing value of **P** even though **E** = 0.

Equations 30-22 and 30-23 show that for isotropic materials, to which a single dielectric constant κ can be assigned, **D** and **P** both point in the direction of **E** at any given point.

The definition of **D** given by Eq. 30-22 allows us to write Eq. 30-15, that is, Gauss's law in the presence of a dielectric, simply as

$$\oint \mathbf{D} \cdot d\mathbf{S} = q, \tag{30-24}$$

where, as before, q represents the free charge only, the induced surface charges being excluded.

▶ **Example 6.** In Figure 30-13, using data from Example 5, calculate E, D, and P: (*a*) in the dielectric and (*b*) in the air gap.

(*a*) The electric field in the dielectric is calculated in Example 5 to be 1.43×10^3 volts/meter. From Eq. 30-22,

$$\begin{aligned} D &= \kappa\epsilon_0 E \\ &= (7.0)(8.9 \times 10^{-12}\ \text{coul}^2/\text{nt-m}^2)(1.43 \times 10^3\ \text{volts/meter}) \\ &= 8.9 \times 10^{-8}\ \text{coul/meter}^2 \end{aligned}$$

and, from Eq. 30-23,

$$\begin{aligned} P &= \epsilon_0(\kappa - 1)E \\ &= (8.9 \times 10^{-12}\ \text{coul}^2/\text{nt-m}^2)(7.0 - 1)(1.43 \times 10^3\ \text{volts/meter}) \\ &= 7.5 \times 10^{-8}\ \text{coul/meter}^2. \end{aligned}$$

(*b*) The electric field E_0 in the air gap is calculated in Example 5 to be 1.00×10^4 volts/meter. From Eq. 30-22,

$$\begin{aligned} D_0 &= \kappa\epsilon_0 E_0 \\ &= (1)(8.9 \times 10^{-12}\ \text{coul}^2/\text{nt-m}^2)(1.00 \times 10^4\ \text{volts/meter}) \\ &= 8.9 \times 10^{-8}\ \text{coul/meter}^2 \end{aligned}$$

and, from Eq. 30-23, recalling that $\kappa = 1$ in the air gap,

$$P_0 = \epsilon_0(\kappa - 1)E_0 = 0.$$

Note that **P** vanishes outside the dielectric, **D** has the same value in the dielectric and in the gap, and **E** has different values in the dielectric and in the gap. The student should verify that Eq. 30-21 ($\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P}$) is correct both in the gap and in the dielectric.

It can be shown from Maxwell's equations that no matter how complex the problem the component of **D** *normal* to the surface of the dielectric has the same value on each side of the surface. In this problem **D** itself is normal to the surface, there being no component but the normal one. It can also be shown that the component of **E** *tangential* to the dielectric surface has the same value on each side of the surface. This *boundary condition*, like the one for **D**, is trivial in this problem, both tangential components being zero. In more complex problems these boundary conditions on **D** and **E** are very important. Table 30-2 summarizes the properties of the electric vectors **E**, **D**, and **P**. ◀

Table 30–2

THREE ELECTRIC VECTORS

Name	Symbol	Associated with	Boundary Condition
Electric field strength	$\mathbf{E}$	All charges	Tangential component continuous
Electric displacement	$\mathbf{D}$	Free charges only	Normal component continuous
Polarization (electric dipole moment per unit volume)	$\mathbf{P}$	Polarization charges only	Vanishes in a vacuum

Defining equation for **E**	$\mathbf{F} = q\mathbf{E}$	Eq. 27–2
General relation among the three vectors	$\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P}$	Eq. 30–21
Gauss's law when dielectric media are present	$\oint \mathbf{D}\cdot d\mathbf{S} = q$ (q = free charge only)	Eq. 30–24
Empirical relations for certain dielectric materials *	$\mathbf{D} = \kappa\epsilon_0\mathbf{E}$ $\mathbf{P} = (\kappa - 1)\epsilon_0\mathbf{E}$	Eq. 30–22 Eq. 30–23

* Generally true, with κ independent of **E**, except for certain materials called *ferroelectrics*; see footnote on page 758.

30–7 Energy Storage in an Electric Field

In Section 29–6 we saw that all charge configurations have a certain *electric potential energy* U, equal to the work W (which may be positive or negative) that must be done to assemble them from their individual components, originally assumed to be infinitely far apart and at rest. This potential energy reminds us of the potential energy stored in a compressed spring or the gravitational potential energy stored in, say, the earth-moon system.

For a simple example, work must be done to separate two equal and opposite charges. This energy is stored in the system and can be recovered if the charges are allowed to come together again. Similarly, a charged capacitor has stored in it an electrical potential energy U equal to the work W required to charge it. This energy can be recovered if the capacitor is allowed to discharge. We can visualize the work of charging by imagining that an external agent pulls electrons from the positive plate and pushes them onto the negative plate, thus bringing about the charge separation; normally the work of charging is done by a battery, at the expense of its store of chemical energy.

Suppose that at a time t a charge $q'(t)$ has been transferred from one plate to the other. The potential difference $V(t)$ between the plates at that moment will be $q'(t)/C$. If an extra increment of charge dq' is transferred, the small amount of additional work needed will be

$$dW = V\,dq = \left(\frac{q'}{C}\right) dq'.$$

If this process is continued until a total charge q has been transferred, the total work will be found from

$$W = \int dW = \int_0^q \frac{q'}{C}\,dq' = \frac{1}{2}\frac{q^2}{C}. \tag{30–25}$$

From the relation $q = CV$ we can also write this as

$$W\ (= U) = \tfrac{1}{2}CV^2. \tag{30–26}$$

It is reasonable to suppose that the energy stored in a capacitor resides in the electric field. As q or V in Eqs. 30–25 and 30–26 increase, for example, so does the electric field E; when q and V are zero, so is E.

In a parallel-plate capacitor, neglecting fringing, the electric field has the same value for all points between the plates. Thus the *energy density u*, which is the stored energy per unit volume, should also be uniform; u (see Eq. 30–26) is given by

$$u = \frac{U}{Ad} = \frac{\frac{1}{2}CV^2}{Ad},$$

where Ad is the volume between the plates. Substituting the relation $C = \kappa\epsilon_0 A/d$ (Eq. 30–8) leads to

$$u = \frac{\kappa\epsilon_0}{2}\left(\frac{V}{d}\right)^2.$$

However, V/d is the electric field strength E, so that

$$u = \tfrac{1}{2}\kappa\epsilon_0 E^2. \tag{30–27}$$

Although this equation was derived for the special case of a parallel-plate capacitor, it is true in general. *If an electric field* **E** *exists at any point in space, we can think of that point as the site of stored energy in amount, per unit volume, of* $\frac{1}{2}\kappa\epsilon_0 E^2$.

▶ **Example 7.** A capacitor C_1 is charged to a potential difference V_0. This charging battery is then removed and the capacitor is connected as in Fig. 30–14 to an uncharged capacitor C_2.

(*a*) What is the final potential difference V across the combination?

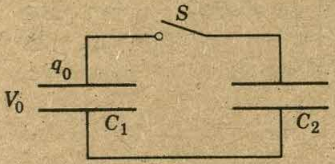


Fig. 30–14 Example 7. C_1 is charged and then connected to C_2 by closing switch S.

The original charge q_0 is now shared by the two capacitors. Thus

$$q_0 = q_1 + q_2.$$

Applying the relation $q = CV$ to each of these terms yields

$$C_1V_0 = C_1V + C_2V$$

or

$$V = V_0 \frac{C_1}{C_1 + C_2}.$$

This suggests a way to measure an unknown capacitance (C_2, say) in terms of a known one.

(*b*) What is the stored energy before and after the switch in Fig. 30–14 is thrown?

The initial stored energy is

$$U_0 = \tfrac{1}{2}C_1V_0^2.$$

The final stored energy is

$$U = \tfrac{1}{2}C_1V^2 + \tfrac{1}{2}C_2V^2 = \tfrac{1}{2}(C_1 + C_2)\left(\frac{V_0C_1}{C_1 + C_2}\right)^2 = \left(\frac{C_1}{C_1 + C_2}\right)U_0.$$

Thus U is less than U_0! The "missing" energy appears as heat in the connecting wires as the charges move through them.

Example 8. A parallel-plate capacitor has plates with area A and separation d. A battery charges the plates to a potential difference V_0. The battery is then disconnected, and a dielectric slab of thickness d is introduced. Calculate the stored energy both before and after the slab is introduced and account for any difference.

The energy U_0 before introducing the slab is

$$U_0 = \tfrac{1}{2}C_0V_0^2.$$

After the slab is in place, we have

$$C = \kappa C_0 \qquad \text{and} \qquad V = V_0/\kappa$$

and thus

$$U = \tfrac{1}{2}CV^2 = \tfrac{1}{2}\kappa C_0\left(\frac{V_0}{\kappa}\right)^2 = \frac{1}{\kappa}U_0.$$

The energy after the slab is introduced is *less* by a factor $1/\kappa$. The "missing" energy would be apparent to the person who inserted the slab. He would feel a "tug" on the slab and would have to restrain it if he wished to insert the slab without acceleration. This means that he would have to do negative work on it, or, alternatively, that the condenser + slab system would do positive work on him. This positive work is

$$W = U_0 - U = \tfrac{1}{2}C_0V_0^2\left(1 - \frac{1}{\kappa}\right).$$

Figure 30-15 shows how the forces that do this work arise, in terms of attraction between the free charge on the plates and the induced surface charges that appear on the slab when it is introduced between the plates.

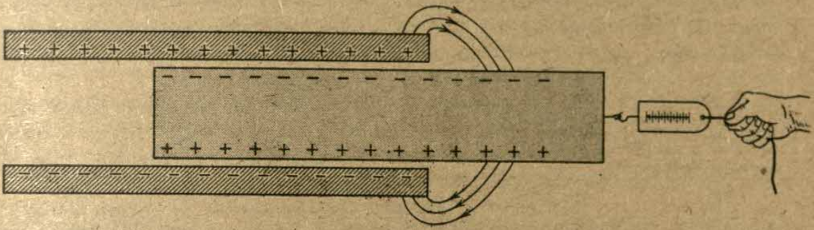


Fig. 30-15 When we introduce a dielectric slab into a charged capacitor, as shown, forces arise which tend to pull the slab into the capacitor.

Example 9. A conducting sphere of radius R, in a vacuum, carries a charge q. (a) Compute the total electrostatic energy stored in the surrounding space.

At any radius r from the center of the sphere (assuming $r > R$) E is given by

$$E = \frac{1}{4\pi\epsilon_0}\frac{q}{r^2}.$$

The energy density at any radius r is found from Eq. 30-27, assuming $\kappa = 1$, or

$$u = \tfrac{1}{2}\epsilon_0 E^2 = \frac{q^2}{32\pi^2\epsilon_0 r^4}.$$

The energy dU that lies in a spherical shell between the radii r and $r + dr$ is

$$dU = (4\pi r^2)(dr)u = \frac{q^2}{8\pi\epsilon_0}\frac{dr}{r^2},$$

where $(4\pi r^2)(dr)$ is the volume of the spherical shell. The total energy U is found by integration, or

$$U = \int dU = \frac{q^2}{8\pi\epsilon_0}\int_R^\infty \frac{dr}{r^2} = \frac{q^2}{8\pi\epsilon_0 R}.$$

Note that this relation follows at once from Eq. 30-25 ($U = q^2/2C$), where C (see p. 651) is the capacitance ($= 4\pi\epsilon_0 R$) of an isolated sphere of radius R.

(b) What is the radius R_0 of a spherical surface such that half the stored energy lies within it?

In the equation just given we put

$$\tfrac{1}{2}U = \frac{q^2}{8\pi\epsilon_0}\int_R^{R_0}\frac{dr}{r^2}$$

or

$$\frac{q^2}{16\pi\epsilon_0 R} = \frac{q^2}{8\pi\epsilon_0}\left(\frac{1}{R} - \frac{1}{R_0}\right),$$

which yields, after some rearrangement,

$$R_0 = 2R.$$

QUESTIONS

1. A capacitor is connected across a battery. (*a*) Why does each plate receive a charge of exactly the same magnitude? (*b*) Is this true even if the plates are of different sizes?

2. Can there be a potential difference between two adjacent conductors that carry the same positive charge?

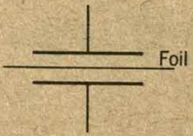


Fig. 30–16

3. The relation $\sigma \propto 1/R$ suggests that the charge placed on an isolated conductor concentrates on points and avoids flat surfaces, where $R = \infty$. How do we reconcile this with Fig. 30–3 in which the charge is definitely on the flat surface of either plate?

4. A sheet of aluminum foil of negligible thickness is placed between the plates of a capacitor as in Fig. 30–16. What effect has it on the capacitance if (*a*) the foil is electrically insulated and (*b*) the foil is connected to the upper plate.

5. Discuss similarities and differences when (*a*) a dielectric slab and (*b*) a conducting slab are inserted between the plates of a parallel-plate capacitor. Assume the slab thicknesses to be one-half the plate separation.

6. An oil-filled parallel-plate capacitor has been designed to have a capacitance C and to operate safely at or below a certain maximum potential difference V_m without arcing over. However, the designer did not do a good job and the capacitor occasionally arcs over. What can be done to redesign the capacitor, keeping C and V_m unchanged and using the same dielectric?

7. Would you expect the dielectric constant, for substances containing permanent molecular electric dipoles, to vary with temperature?

8. What is your estimate of the amount by which the center of positive charge and the center of negative charge are displaced in a situation like that of Fig. 30–9*b*? An intuitive guess is all that is called for.

9. For a given potential difference does a capacitor store more or less charge with a dielectric than it does without a dielectric (vacuum)? Explain in terms of the microscopic picture of the situation.

10. An isolated conducting sphere is given a positive charge. Does its mass increase, decrease, or remain the same?

11. A dielectric slab is inserted in one end of a charged parallel-plate capacitor (the plates being horizontal and the charging battery having been disconnected) and then released. Describe what happens. Neglect friction.

12. A capacitor is charged by using a battery, which is then disconnected. A dielectric slab is then slipped between the plates. Describe qualitatively what happens to the charge, the capacitance, the potential difference, the electric field strength, and the stored energy.

13. While a capacitor remains connected to a battery, a dielectric slab is slipped between the plates. Describe qualitatively what happens to the charge, the capacitance, the potential difference, the electric field strength, and the stored energy. Is work required to insert the slab?

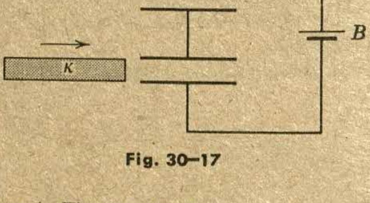


Fig. 30–17

14. Two identical capacitors are connected as shown in Fig. 30–17. A dielectric slab is slipped between the plates of one capacitor, the battery remaining connected. Describe qualitatively what happens to the charge, the capacitance, the potential difference, the electric field strength, and the stored energy for each capacitor.

15. Show that the dielectric constant of a conductor can be taken to be infinitely great.

PROBLEMS

1. A potential difference of 300 volts is applied to a 2.0-μf capacitor and an 8.0-μf capacitor connected in series. (*a*) What are the charge and the potential difference for each capacitor? (*b*) The charged capacitors are reconnected with their positive plates together and their negative plates together, no external voltage being applied. What are the charge and the potential difference for each? (*c*) The charged capacitors in (*a*) are reconnected with plates of *opposite* sign together. What are the charge and the potential difference for each?

2. Calculate the capacitance of the earth, viewed as a spherical conductor of radius 6400 km.

3. A 100-$\mu\mu$f capacitor is charged to a potential difference of 50 volts, the charging battery then being disconnected. The capacitor is then connected, as in Fig. 30–14, to a second capacitor. If the measured potential difference drops to 35 volts, what is the capacitance of this second capacitor?

4. If we solve Eq. 30–7 for ϵ_0, we see that its mks units are farads/meter. Show that these units are equivalent to those obtained earlier for ϵ_0, namely coul^2/nt-m^2.

5. Figure 30–18 shows two capacitors in series, the rigid center section of length b being movable vertically. Show that the equivalent capacitance of the series combination is independent of the position of the center section and is given by

$$C = \frac{\epsilon_0 A}{a - b}.$$

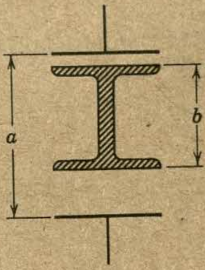


Fig. 30–18

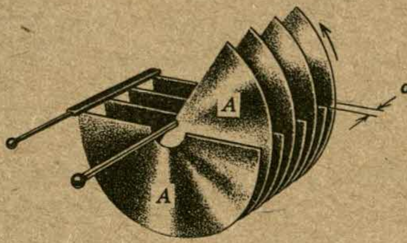


Fig. 30–19

6. In Fig. 30–19 a variable air capacitor of the type used in tuning radios is shown. Alternate plates are connected together, one group being fixed in position, the other group being capable of rotation. Consider a pile of n plates of alternate polarity, each having an area A and separated from adjacent plates by a distance d. Show that this capacitor has a maximum capacitance of

$$C = \frac{(n - 1)\epsilon_0 A}{d}.$$

7. A spherical capacitor consists of two concentric spherical shells of radii a and b, with $b > a$. Show that its capacitance is

$$C = 4\pi\epsilon_0 \frac{ab}{b - a}.$$

8. A capacitor has square plates, each of side a, making an angle of θ with each other as shown in Fig. 30–20. Show that for small θ the capacitance is given by

$$C = \frac{\epsilon_0 a^2}{d}\left(1 - \frac{a\theta}{2d}\right)$$

(Hint: The capacitor may be divided into differential strips which are effectively in parallel.)

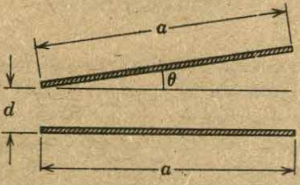


Fig. 30–20

9. Suppose that the two spherical shells of a spherical capacitor have their radii approximately equal. Under these conditions the device approximates a parallel plate capacitor with $b - a = d$. Show that the formula in Problem 7 does indeed reduce to Eq. 30–7 in this case.

10. A parallel-plate capacitor has circular plates of 8.0-cm radius and 1.0-mm separation. What charge will appear on the plates if a potential difference of 100 volts is applied?

11. In Fig. 30–21 find the equivalent capacitance of the combination. Assume that $C_1 = 10\ \mu f$, $C_2 = 5\ \mu f$, $C_3 = 4\ \mu f$, and $V = 100$ volts.

12. In Fig. 30–21 suppose that capacitor C_3 breaks down electrically, becoming equivalent to a conducting path. What *changes* in (*a*) the charge and (*b*) the potential difference occur for capacitor C_1?

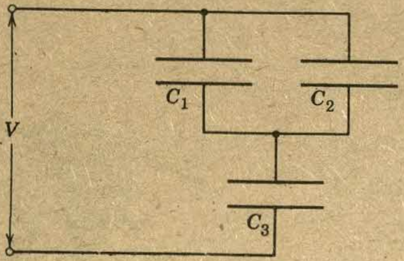


Fig. 30–21

13. Find the effective capacitance between points x and y in Fig. 30–22. Assume that $C_2 = 10\ \mu f$ and that the other capacitors are all $4.0\ \mu f$. (Hint: Apply a potential difference V between x and y and write down all the relationships that involve the charges and potential differences for the separate capacitors.)

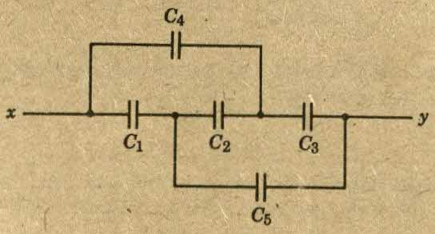


Fig. 30–22

14. If you have available several 2.0-μf capacitors, each capable of withstanding 200 volts without breakdown, how would you assemble a combination having an equivalent capacitance of (*a*) 0.40 μf or of (*b*) 1.2 μf, each capable of withstanding 1000 volts?

15. In Fig. 30–23 find the equivalent capacitance of the combination. Assume that $C_1 = 10\ \mu\text{f}$, $C_2 = 5\ \mu\text{f}$, $C_3 = 4\ \mu\text{f}$, and $V = 100$ volts.

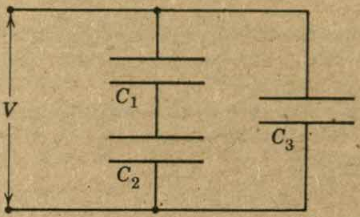


Fig. 30–23

16. In Fig. 30–24 the battery B supplies 12 volts. (*a*) Find the charge on each capacitor when switch S_1 is closed and (*b*) when switch S_2 is also closed. Take $C_1 = 1\ \mu\text{f}$, $C_2 = 2\ \mu\text{f}$, $C_3 = 3\ \mu\text{f}$, and $C_4 = 4\ \mu\text{f}$.

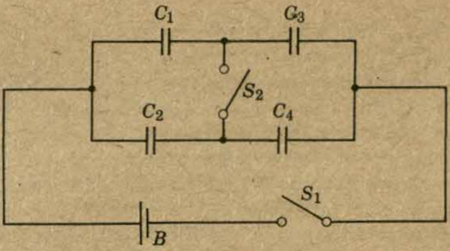


Fig. 30–24

17. A parallel-plate capacitor is filled with two dielectrics as in Fig. 30–25. Show that the capacitance is given by

$$C = \frac{\epsilon_0 A}{d}\left(\frac{\kappa_1 + \kappa_2}{2}\right).$$

Check this formula for all the limiting cases that you can think of. (Hint: Can you justify regarding this arrangement as two capacitors in parallel?)

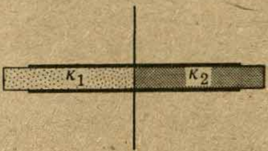


Fig. 30–25

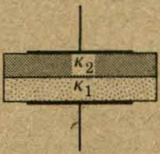


Fig. 30–26

18. A parallel-plate capacitor is filled with two dielectrics as in Fig. 30–26. Show that the capacitance is given by

$$C = \frac{2\epsilon_0 A}{d}\left(\frac{\kappa_1\kappa_2}{\kappa_1 + \kappa_2}\right).$$

Check this formula for all the limiting cases that you can think of. (Hint: Can you justify regarding this arrangement as two capacitors in series?)

19. A dielectric slab of thickness b is inserted between the plates of a parallel-plate capacitor of plate separation d. Show that the capacitance is given by

$$C = \frac{\kappa\epsilon_0 A}{\kappa d - b(\kappa - 1)}.$$

(Hint: Derive the formula following the pattern of Example 5.) Does this formula predict the correct numerical result of Example 5? Does the formula seem reasonable for the special cases of $b = 0$, $\kappa = 1$, and $b = d$?

20. A slab of copper of thickness b is thrust into a parallel-plate capacitor as shown in Fig. 30–27; it is exactly halfway between the plates. What is the capacitance before and after the slab is introduced?

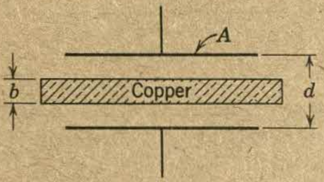


Fig. 30–27

21. For making a capacitor you have available two plates of copper, a sheet of mica (thickness = 0.10 mm, $\kappa = 6$), a sheet of glass (thickness = 2.0 mm, $\kappa = 7$), and a slab of paraffin (thickness = 1.0 cm, $\kappa = 2$). To obtain the largest capacitance, which sheet (or sheets) should you place between the copper plates?

22. In Example 5, suppose that the 100-volt battery remains connected during the time that the dielectric slab is being introduced. Calculate (*a*) the charge on the capacitor plates, (*b*) the electric field in the gap, (*c*) the electric field in the slab, and (*d*) the capacitance. For all of these quantities give the numerical values before and after the slab is introduced. Contrast your results with those of Example 5 by constructing a tabular listing.

23. A parallel-plate capacitor has a capacitance of 100 $\mu\mu$f, a plate area of 100 cm^2, and a mica dielectric. At 50 volts potential difference, calculate (*a*) E in the mica, (*b*) the free charge on the plates, and (*c*) the induced surface charge.

24. Two parallel plates of area 100 cm^2 are each given equal but opposite charges of 8.9×10^{-7} coul. Within the dielectric material filling the space between the plates the electric field strength is 1.4×10^6 volts/meter. (*a*) Find the dielectric constant of the material. (*b*) Determine the magitude of the charge induced on each dielectric surface.

25. Hard rubber has a dielectric constant of 2.8 and a dielectric strength of 18×10^6 volts/meter. If it is used as the dielectric material in a parallel-plate capacitor, what minimum area may the plates of the capacitor have in order that the capacitance be 7.0×10^{-2} μf and that the capacitor be able to withstand a potential difference of 4000 volts?

26. A cylindrical capacitor has radii a and b as in Fig. 30–4 Show that half the stored electric potential energy lies within a cylinder whose radius is

$$r = \sqrt{ab}.$$

27 An isolated metal sphere whose diameter is 10 cm has a potential of 8000 volts. What is the energy density at the surface of the sphere?

28. A parallel-plate capacitor has plates of area A and separation d and is charged to a potential difference V. The charging battery is then disconnected and the plates are pulled

apart until their separation is $2d$. Derive expressions in terms of A, d, and V for (*a*) the new potential difference, (*b*) the initial and the final stored energy, and (*c*) the work required to separate the plates.

29. Show that the plates of a parallel-plate capacitor attract each other with a force given by

$$F = \frac{q^2}{2\epsilon_0 A}.$$

Prove this by calculating the work necessary to increase the plate separation from x to $x + dx$.

30. In the capacitor of Example 5 the dielectric slab fills half the space between the plates. (*a*) What per cent of the energy is stored in the air gaps? (*b*) What per cent is stored in the slab?

31. A parallel-plate air capacitor has a capacitance of 100 $\mu\mu$f. (*a*) What is the stored energy if the applied potential difference is 50 volts? (*b*) Can you calculate the energy density for points between the plates?

32. For the capacitors of Problem 1, compute the energy stored for the three different connections of parts (*a*), (*b*), and (*c*). Compare your answers and explain any differences.

33. A charge q is placed on the surface of an originally uncharged soap bubble of radius R_0. Because of the mutual repulsion of the charged surface, the radius is increased to a somewhat larger value R. Show that

$$q = [\tfrac{32}{3}\pi^2\epsilon_0 p R_0 R(R^2 + R_0 R + R_0{}^2)]^{1/2}$$

in which p is the pressure of the atmosphere. Find q for $p = 1.00$ atm, $R_0 = 2.00$ cm, and $R = 2.10$ cm. (Hint: The work done by the bubble in pushing back the atmosphere must equal the decrease in the stored electric field energy that accompanies the expansion, from the conservation of energy principle.)

34. Two capacitors (2.0 μf and 4.0 μf) are connected in parallel across a 300-volt potential difference. Calculate the total stored energy in the system.

35. A parallel-connected bank of 2000 5.0-μf capacitors is used to store electric energy. What does it cost to charge this bank to 50,000 volts, assuming a rate of 2¢/kw-hr?

36. In Fig. 30–21 find (*a*) the charge, (*b*) the potential difference, and (*c*) the stored energy for each capacitor. Assume the numerical values of Problem 11.

37. In Fig. 30–23 find (*a*) the charge, (*b*) the potential difference, and (*c*) the stored energy for each capacitor. Assume the numerical values of Problem 15.

Current and Resistance

CHAPTER 31

31-1 Current and Current Density

The free electrons in an isolated metallic conductor, such as a length of copper wire, are in random motion like the molecules of a gas confined to a container. They have no net directed motion along the wire. If a hypothetical plane is passed through the wire, the rate at which electrons pass through it from right to left is the same as the rate at which they pass through from left to right; the *net* rate is zero.

If the ends of the wire are connected to a battery, an electric field will be set up at every point within the wire. If the potential difference maintained by the battery is 10 volts and if the wire (assumed uniform) is 5 meters long, the strength of this field at every point will be 2 volts/meter. This field **E** will act on the electrons and will give them a resultant motion in the direction of $-\mathbf{E}$. We say that an *electric current* i is established; if a net charge q passes through any cross section of the conductor in time t the current, assumed constant, is

$$i = q/t. \tag{31-1}$$

The appropriate mks units are amperes for i, coulombs for q, and seconds for t. The student will recall (Section 26–4) that Eq. 31–1 is the defining equation for the coulomb and that we have not yet given an operational definition of the ampere; we do so in Section 34–4.

If the rate of flow of charge with time is not constant, the current varies with time and is given by the differential limit of Eq. 31–1, or

$$i = dq/dt. \tag{31-2}$$

In the rest of this chapter we consider only constant currents.

The current i is the same for all cross sections of a conductor, even though the cross-sectional area may be different at different points. In the same way the rate at which water (assumed incompressible) flows past any cross section of a pipe is the same even if the cross section varies. The water flows faster where the pipe is smaller and slower where it is larger, so that the volume rate, measured perhaps in gal/min, remains unchanged. This constancy of the electric current follows because charge must be conserved; it does not pile up steadily or drain away steadily from any point in the conductor under the assumed steady-state conditions. In the language of Section 18-3 there are no "sources" or "sinks" of charge.

The existence of an electric field inside a conductor does not contradict Section 28-4, in which we asserted that **E** equals zero inside a conductor. In that section, which dealt with a state in which all net motion of charge had stopped (*electrostatics*), we assumed that the conductor was insulated and that no potential difference was deliberately maintained between any two points on it, as by a battery. In this chapter, which deals with charges in motion, we relax this restriction.

The electric field that acts on the electrons in a conductor does not produce a *net* acceleration because the electrons keep colliding with the atoms (strictly, ions) that make up the conductor. This array of ions, coupled together by strong spring-like forces of electric origin, is called the *lattice* (see Fig. 21-5). The over-all effect of these collisions is to transfer kinetic energy from the accelerating electrons into vibrational energy of the lattice. The electrons acquire a constant average *drift speed* v_d in the direction $-\mathbf{E}$. The analogy is to a marble rolling down a long flight of stairs and not to a marble falling freely from the same height. In the first case the acceleration caused by the (gravitational) field is effectively canceled by the decelerating effects of collisions with the stair treads so that, under proper conditions, the marble rolls down the stairs with zero average acceleration, that is, at constant average speed.

Although in metals the charge carriers are electrons, in electrolytes or in gaseous conductors they may also be positive or negative ions or both. A convention for labeling the directions of currents is needed because charges of opposite sign move in opposite directions in a given field. A positive charge moving in one direction is equivalent in nearly all external effects to a negative charge moving in the opposite direction. Hence, for simplicity and algebraic consistency, *we assume that all charge carriers are positive and we draw the current arrows in the direction that such charges would move.* If the charge carriers are negative, they simply move opposite to the direction of the current arrow (see Fig. 31-1). When we encounter a case (as in the *Hall effect;* see Section 33-5) in which the sign of the charge carriers makes a difference in the external effects, we will disregard the convention and take the actual situation into account.

Current i is a characteristic of a particular conductor. It is a macroscopic quantity, like the mass of an object, the volume of an object, or the length of a rod. A related microscopic quantity is the current density **j**. It is a

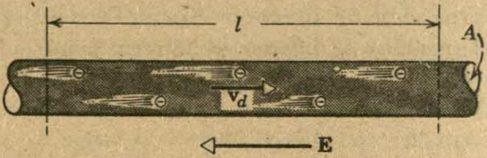


Fig. 31–1 Electrons drift in a direction opposite to the electric field in a conductor.

vector and is characteristic of a point inside a conductor rather than of the conductor as a whole. If the current is distributed uniformly across a conductor of cross-sectional area A, the magnitude of the current density for all points on that cross section is

$$j = i/A. \tag{31-3}$$

The vector $\mathbf{j}$ at any point is oriented in the direction that a positive charge carrier would move at that point. An electron at that point would move in the direction $-\mathbf{j}$.

The general relationship between $\mathbf{j}$ and i is that, for a particular surface in a conductor, i is the flux of the vector $\mathbf{j}$ over that surface, or

$$i = \int \mathbf{j} \cdot d\mathbf{S}, \tag{31-4}$$

where $d\mathbf{S}$ is an element of surface area and the integral is taken over the surface in question. Equation 31–3 (written as $i = jA$) is a special case of this relationship in which the surface of integration is a cross section of the conductor and in which $\mathbf{j}$ is constant over this surface and at right angles to it. However, Eq. 31–4 may be applied to any surface through which we wish to know the current. Equation 31–4 shows clearly that i is a scalar because the integrand $\mathbf{j} \cdot d\mathbf{S}$ is a scalar.

The arrow often associated with the current in a wire does not indicate that current is a vector but merely shows the *sense* of charge flow. Positive charge carriers either move in a certain direction along the wire or in the opposite direction, these two possibilities being represented by + or − in algebraic equations. Note that (*a*) the current in a wire remains unchanged if the wire is bent, tied into a knot, or otherwise distorted, and (*b*) the arrows representing the sense of currents do not in any way obey the laws of vector addition.

The drift speed v_d of charge carriers in a conductor can be computed from the current density j. Figure 31–1 shows the conduction electrons in a wire moving to the right at an assumed constant drift speed v_d. The number of conduction electrons in the wire is nAl where n is the number of conduction electrons per unit volume and Al is the volume of the wire. A charge of magnitude

$$q = (nAl)e$$

passes out of the wire, through its right end, in a time t given by

$$t = \frac{l}{v_d}.$$

The current i is given by

$$i = \frac{q}{t} = \frac{nAle}{l/v_d} = nAev_d.$$

Solving for v_d and recalling that $j = i/A$ (Eq. 31-3) yields

$$v_d = \frac{i}{nAe} = \frac{j}{ne}. \tag{31-5}$$

▶ **Example 1.** An aluminum wire whose diameter is 0.10 in. is welded end to end to a copper wire with a diameter of 0.064 in. The composite wire carries a steady current of 10 amp. What is the current density in each wire?

The current is distributed uniformly over the cross section of each conductor, except near the junction, which means that the current density may be taken as constant for all points within each wire. The cross-sectional area of the aluminum wire is 0.0079 in.^{2} Thus, from Eq. 31-3,

$$j_{\text{Al}} = \frac{i}{A} = \frac{10 \text{ amp}}{0.0079 \text{ in.}^2} = 1300 \text{ amp/in.}^2$$

The cross-sectional area of the copper wire is 0.0032 in.^{2} Thus

$$j_{\text{Cu}} = \frac{i}{A} = \frac{10 \text{ amp}}{0.0032 \text{ in.}^2} = 3100 \text{ amp/in.}^2$$

The fact that the wires are of different materials does not enter into consideration here.

Example 2. What is v_d for the copper wire in Example 1?

We can write the current density for the copper wire as 480 amp/cm^2. To compute n we start from the fact that there is one free electron per atom in copper. The number of atoms per unit volume is dN_0/M where d is the density, N_0 is Avogadro's number, and M is the atomic weight. The number of free electrons per unit volume is then

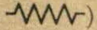
$$n = \frac{dN_0}{M} = \frac{(9.0 \text{ gm/cm}^3)(6.0 \times 10^{23} \text{ atoms/mole})(1 \text{ electron/atom})}{64 \text{ gm/mole}}$$
$$= 8.4 \times 10^{22} \text{ electrons/cm}^3.$$

Finally, v_d is, from Eq. 31-5,

$$v_d = \frac{j}{ne} = \frac{480 \text{ amp/cm}^2}{(8.4 \times 10^{22} \text{ electrons/cm}^3)(1.6 \times 10^{-19} \text{ coul/electron})}$$
$$= 3.6 \times 10^{-2} \text{ cm/sec}.$$

It takes 28 sec for the electrons in this wire to drift 1.0 cm. Would you have guessed that v_d was so low? The drift speed of electrons must not be confused with the speed at which changes in the electric field configuration travel along wires, a speed which approaches that of light. When pressure is applied to one end of a long water-filled tube, a *pressure wave* travels rapidly along the tube. The speed at which *water* moves through the tube is much lower, however. ◀

31–2 Resistance, Resistivity, and Conductivity

If the same potential difference is applied between the ends of a rod of copper and of a rod of wood, very different currents result. The characteristic of the conductor that enters here is its *resistance*. We define the resistance of a conductor (often called a *resistor;* symbol ) between two points by applying a potential difference V between those points, measuring the current i, and dividing:

$$R = V/i. \tag{31-6}$$

If V is in volts and i in amperes, the resistance R will be in *ohms*.

The flow of charge through a conductor is often compared with the flow of water through a pipe, which occurs because there is a difference in pressure between the ends of the pipe, established perhaps by a pump. This pressure difference can be compared with the potential difference established between the ends of a resistor by a battery. The flow of water (ft^3/sec, say) is compared with the current (coul/sec or amp). The rate of flow of water for a given pressure difference is determined by the nature of the pipe. Is it long or short? Is it narrow or wide? Is it empty or filled, perhaps with gravel? These characteristics of the pipe are analogous to the resistance of a conductor.

Primary standards of resistance, kept at the National Bureau of Standards, are spools of wire whose resistances have been accurately measured. Because resistance varies with temperature, these standards, when used, are placed in an oil bath at a controlled temperature. They are made of a special alloy, called *manganin*, for which the change of resistance with temperature is very small. They are carefully annealed to eliminate strains, which also affect the resistance. These primary standard resistors are used chiefly to calibrate secondary standards for other laboratories.

Operationally, the primary resistance standards are not measured by using Eq. 31–6 but are measured in an indirect way which involves magnetic fields. Equation 31–6 is, in fact, used to measure V, by setting up an accurately known current i (using a *current balance;* see Section 34–4) in an accurately known resistance R. This operational procedure for potential difference is the one normally used in place of the conceptual definition introduced in Section 29–1, in which one measures the work per unit charge required to move a test charge between two points.

Related to resistance is the *resistivity* ρ, which is a characteristic of a material rather than of a particular specimen of a material; it is defined, for isotropic materials,* from

$$\rho = \frac{E}{j}. \tag{31-7}$$

The resistivity of copper is 1.7×10^{-8} ohm-m; that of fused quartz is about 10^{16} ohm-m. Few physical properties are measurable over such a range of values; Table 31–1 lists some values for common metals.

---

* These are materials whose properties (electrical in this case) do not vary with direction in the material.

Table 31-1

PROPERTIES OF METALS AS CONDUCTORS

	Resistivity (at 20°C), ohm-m	Temperature Coefficient of Resistivity,* α, per C°	Density, gm/cm^3	Melting Point, °C
Aluminum	2.8×10^{-8}	3.9×10^{-3}	2.7	659
Copper	1.7×10^{-8}	3.9×10^{-3}	8.9	1080
Carbon (amorphous)	3.5×10^{-5}	-5×10^{-4}	1.9	3500
Iron	1.0×10^{-7}	5.0×10^{-3}	7.8	1530
Manganin	4.4×10^{-7}	1×10^{-5}	8.4	910
Nickel	6.8×10^{-8}	6×10^{-3}	8.9	1450
Silver	1.6×10^{-8}	3.8×10^{-3}	10.5	960
Steel	1.8×10^{-7}	3×10^{-3}	7.7	1510
Wolfram (tungsten)	5.6×10^{-8}	4.5×10^{-3}	19	3400

* This quantity, defined from

$$\alpha = \frac{1}{\rho}\frac{d\rho}{dT} \tag{31-8}$$

is the fractional change in resistivity ($d\rho/\rho$) per unit change in temperature. It varies with temperature, the values here referring to 20°C. For copper ($\alpha = 3.9 \times 10^{-3}/\text{C}^\circ$) the resistivity increases by 0.39 per cent for a temperature increase of 1 C° near 20°C. Note that α for carbon is negative, which means that the resistivity *decreases* with increasing temperature.

Consider a cylindrical conductor, of cross-sectional area A and length l, carrying a steady current i. Let us apply a potential difference V between its ends. If the cylinder cross sections at each end are equipotential surfaces, the electric field strength and the current density will be constant for all points in the cylinder and will have the values

$$E = \frac{V}{l} \quad \text{and} \quad j = \frac{i}{A}.$$

The resistivity ρ may then be written as

$$\rho = \frac{E}{j} = \frac{V/l}{i/A}.$$

But V/i is the resistance R which leads to

$$R = \rho\frac{l}{A}. \tag{31-9}$$

V, i, and R are *macroscopic* quantities, applying to a particular body or extended region. The corresponding *microscopic* quantities are $\mathbf{E}$, $\mathbf{j}$, and ρ; they have values at every point in a body. The macroscopic quantities are related to each other by Eq. 31–6 ($V = iR$) and the microscopic quantities by Eq. 31–7, which can be written in vector form as $\mathbf{E} = \mathbf{j}\rho$.

The macroscopic quantities can be found by integrating over the microscopic quantities, using relations already given, namely

$$i = \int \mathbf{j} \cdot d\mathbf{S} \tag{31–4}$$

and

$$V_{ab} = -\int_a^b \mathbf{E} \cdot d\mathbf{l}. \tag{29–5}$$

The integral in Eq. 31–4 is a surface integral, carried out over any cross section of the conductor. The integral in Eq. 29–5 is a line integral carried out along an arbitrary line drawn along the conductor, connecting any two equipotential surfaces, identified by a and b. For a long wire connected to a battery equipotential surface a might be chosen as a cross section of the wire near the positive battery terminal and b might be a cross section near the negative terminal.

The resistance of a conductor between a and b can be expressed in microscopic terms by dividing the two equations, or

$$R = \frac{V_{ab}}{i} = \frac{-\int_a^b \mathbf{E} \cdot d\mathbf{l}}{\int \mathbf{j} \cdot d\mathbf{S}}.$$

If the conductor is a long cylinder of cross section A and length l and if points a and b are its ends, the foregoing equation for R (see Eq. 31–7) reduces to

$$R = \frac{El}{jA} = \rho \frac{l}{A},$$

which is Eq. 31–9.

The macroscopic quantities V, i, and R are of primary interest when we are making electrical measurements on real conducting objects. They are the quantities that one reads on meters. The microscopic quantities $\mathbf{E}$, $\mathbf{j}$, and ρ are of primary importance when we are concerned with the fundamental behavior of matter (rather than of specimens of matter), as we usually are in the research area of *solid state physics*. Thus Section 31–4 deals appropriately with an atomic view of the *resistivity* of a metal and not of the *resistance* of a metallic specimen. The microscopic quantities are also important when we are interested in the interior behavior of irregularly shaped conducting objects.

▶ **Example 3.** A rectangular carbon block has dimensions 1.0 cm × 1.0 cm × 50 cm. (*a*) What is the resistance measured between the two square ends? (*b*) Between two opposing rectangular faces? The resistivity of carbon at 20°C is 3.5×10^{-5} ohm-m.

(*a*) The area of a square end is 1.0 cm^2 or 1.0×10^{-4} meter^2. Equation 31–9 gives for the resistance between the square ends:

$$R = \rho \frac{l}{A} = \frac{(3.5 \times 10^{-5}\ \text{ohm-m})(0.50\ \text{meter})}{1.0 \times 10^{-4}\ \text{meter}^2} = 0.18\ \text{ohm}.$$

(*b*) For the resistance between opposing rectangular faces (area = 5.0×10^{-3} meter^2), we have

$$R = \rho \frac{l}{A} = \frac{(3.5 \times 10^{-5}\ \text{ohm-m})(10^{-2}\ \text{meter})}{5.0 \times 10^{-3}\ \text{meter}^2} = 7.0 \times 10^{-5}\ \text{ohm}.$$

Thus a given conductor can have a number of resistances, depending on how the potential difference is applied to it. The ratio of resistances for these two cases is 2600. We assume in each that the potential difference is applied to the block in such a way that the surfaces between which the resistance is desired are equipotential. Otherwise Eq. 31–9 would not be valid. ◀

Figure 31–2 shows (solid curve) how the resistivity of copper varies with temperature. Sometimes, for practical use, such data are expressed in equation form. If we are interested in only a limited range of temperatures extending, say, from 0 to 500°C, we can fit a straight line to the curve of Fig. 31–2, making it pass through two arbitrarily selected points; see the dashed line. We choose the point labeled T_0, ρ_0 in the figure as a reference point, T_0 being 0°C in this case and ρ_0 being 1.56×10^{-8} ohm-m. The resistivity ρ at any temperature T can be found from the empirical equation of the dashed straight line in Fig. 31–2, which is

$$\rho = \rho_0[1 + \bar{\alpha}(T - T_0)]. \tag{31–10}$$

This relation shows correctly that $\rho \to \rho_0$ as $T \to T_0$.

If we solve Eq. 31–10 for $\bar{\alpha}$, we obtain

$$\bar{\alpha} = \frac{1}{\rho_0}\frac{\rho - \rho_0}{T - T_0}.$$

Comparison with Eq. 31–8 shows that $\bar{\alpha}$ is a *mean temperature coefficient of resistivity* for a selected pair of temperatures rather than the temperature coefficient of resistivity at a particular temperature, which is the definition of α. For most practical purposes Eq. 31–10 gives results that are within the acceptable range of accuracy.

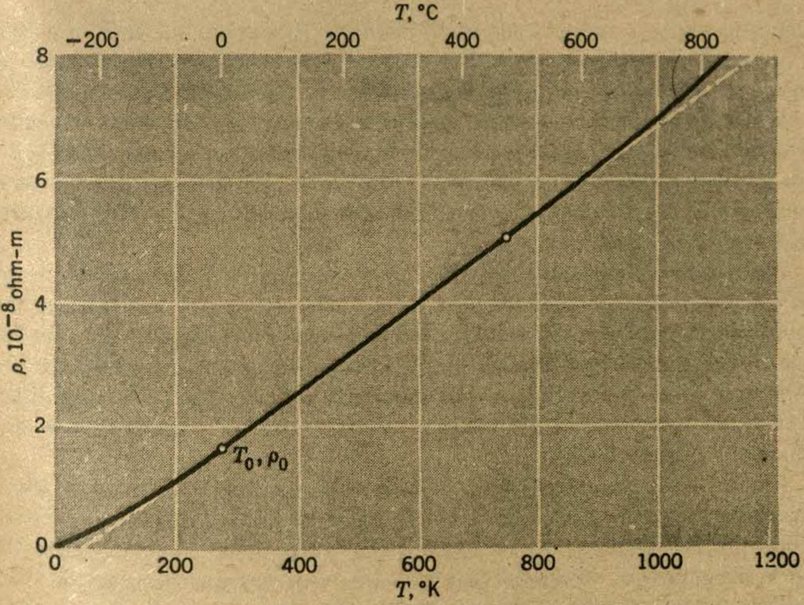


Fig. 31–2 The resistivity of copper as a function of temperature. The dashed line is an approximation chosen to fit the curve at the two circled points. The point marked T_0, ρ_0 is chosen as a reference point.

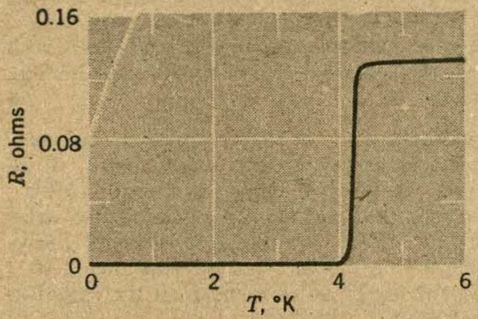


Fig. 31–3 The resistance of mercury disappears below about 4°K.

The curve of Fig. 31–2 does not go to zero at the absolute zero of temperature, even though it appears to do so, the residual resistivity at this temperature being 0.02×10^{-8} ohm-m. For many substances the resistance *does* become zero at some low temperature. Figure 31–3 shows the resistance of a specimen of mercury for temperatures below 6°K. In the space of about 0.05 K° the resistance drops abruptly to an immeasurably low value. This phenomenon, called *superconductivity*,* was discovered by Kamerlingh Onnes in the Netherlands in 1911. The resistance of materials in the superconducting state seems to be truly zero; currents, once established in closed superconducting circuits, persist for weeks without diminution, even though there is no battery in the circuit. If the temperature is raised slightly above the superconducting point, such currents drop rapidly to zero.

31–3 Ohm's Law

Let us apply a variable potential difference V between the ends of a 100-foot coil of #18 copper wire. For each applied potential difference, let us measure the current i and plot it against V as in Fig. 31–4. The straight line that results means that *the resistance of this conductor is the same no matter what applied voltage is used to measure it.* This important result, which holds for metallic conductors, is known as *Ohm's law.* We assume that the temperature of the conductor is essentially constant throughout the measurements.

Many conductors do not obey Ohm's law. Figure 31–5, for example, shows a V–i plot for a type 2A3 vacuum tube. The plot is not straight and the resistance depends on the voltage used to measure it. Also, the current for this device is almost vanishingly small if the polarity of the applied potential difference is reversed. For metallic conductors the current reverses direction when the potential difference is reversed, but its magnitude does not change.

Figure 31–6 shows a typical V–i plot for another nonohmic device, a *thermistor*. This is a semiconductor (see p. 650) with a large and negative temperature coefficient of resistivity α (see Table 31–1) that varies greatly with temperature. We note that two different currents through the thermistor can correspond to the same potential difference between its ends.

---

* See "Superconductivity" by B. T. Matthias, *Scientific American*, p. 92, November 1957.

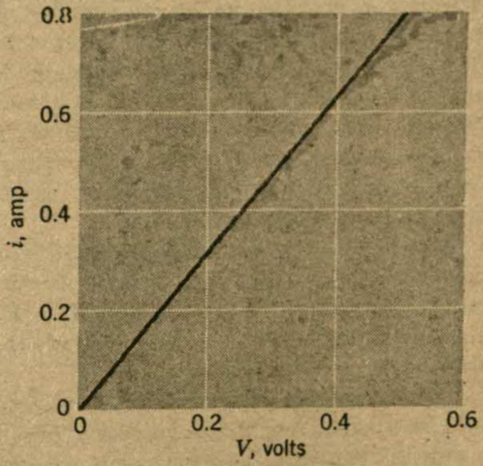


Fig. 31-4 The current in a particular copper conductor as a function of potential difference. This conductor obeys Ohm's law.

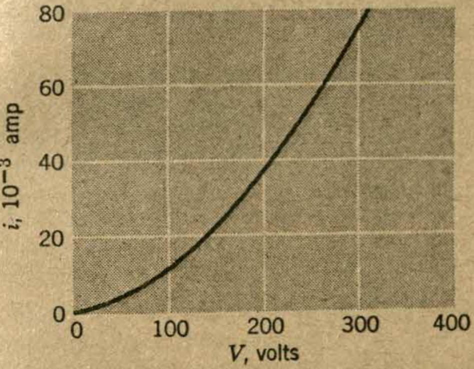
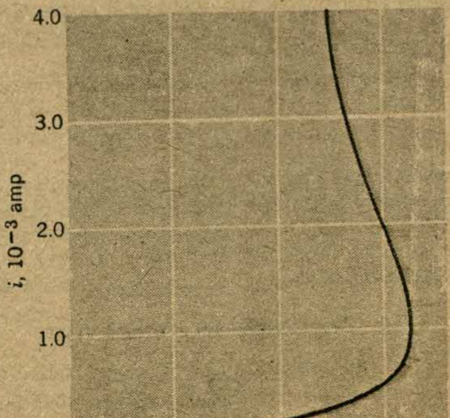


Fig. 31-5 The current in a type 2A3 vacuum tube as a function of potential difference. This conductor does not obey Ohm's law.

Fig. 31-6 A plot of current as a function of potential difference in a Western Electric 1-B thermistor. The curve shows how the voltage across the thermistor varies as the current through it is increased. The shape of the curve can be accounted for in terms of the large negative temperature coefficient of resistivity of the material of which the device is made.



Thermistors are often used to measure the rate of energy flow in microwave beams by allowing the microwave beam to fall on the thermistor and heat it. The relatively small temperature rise so produced results in a relatively large change in resistance, which serves as a measure of the microwave power. Modern electronics, and therefore much of the character of our present technological civilization, depends in a fundamental way on the fact that many conductors, such as vacuum tubes, crystal rectifiers, thermistors, and transistors, do *not* obey Ohm's law.

We stress that the relationship $V = iR$ is *not* a statement of Ohm's law. A conductor obeys this law only if its $V - i$ curve is linear, that is, if R is independent of V and i. The relationship $R = V/i$ remains as the general definition of the resistance of a conductor whether or not the conductor obeys Ohm's law.

The microscopic equivalent of the relationship $V = iR$ is Eq. 31–7, or $\mathbf{E} = \mathbf{j}\rho$. A conducting *material* is said to obey Ohm's law if a plot of E versus j is linear, that is, if the resistivity ρ is independent of E and j. Ohm's law is a specific property of certain materials and is not a general law of electromagnetism, for example, like Gauss's law.

A close analogy exists between the flow of charge because of a potential difference and the flow of heat because of a temperature difference. Consider a thin electrically conducting slab of thickness Δx and area A. Let a potential difference ΔV be maintained between opposing faces. The current i is given by Eqs. 31–6 ($i = V/R$) and 31–9 ($R = \rho l/A$), or

$$i = \frac{\Delta V}{R} = \frac{\Delta VA}{\rho \Delta x}.$$

In the limiting case of a slab of thickness dx this becomes

$$i = \frac{1}{\rho} A \frac{dV}{dx}$$

or

$$\frac{dq}{dt} = -\sigma A \frac{dV}{dx}, \tag{31–11}$$

where σ ($= 1/\rho$) is the *conductivity* of the material. Since positive charge flows in the direction of decreasing V, we introduce a minus sign into Eq. 31–11, that is, dq/dt is positive when dV/dx is negative.

The analogous heat flow equation (see Section 22–4) is

$$\frac{dQ}{dt} = -kA \frac{dT}{dx}, \tag{31–12}$$

which shows that k, the thermal conductivity, corresponds to σ and dT/dx, the temperature gradient, corresponds to dV/dx, the potential gradient. There is more than a formal mathematical analogy between Eqs. 31–11 and 31–12. Both heat energy and charge are carried by the free electrons in a metal; empirically, a good electrical conductor (silver, say) is also a good heat conductor and conversely.

31–4 Resistivity—an Atomic View

One can understand why metals obey Ohm's law on the basis of simple classical ideas. If these ideas are modified when necessary by the requirements of quantum physics, it is possible to go further and to calculate theoretical values of the resistivity ρ for various metals. These calculations are not simple, but when they have been carried out the agreement with the experimental value of ρ has usually been good.

In a metal the valence electrons are not attached to individual atoms but are free to move about within the lattice and are called *conduction electrons*. In copper there is one such electron per atom, the other 28 remaining bound to the copper nuclei to form ionic cores.

The speed distribution of conduction electrons can be described correctly only in terms of quantum physics. For our purposes, however, it suffices to consider only a suitably defined average speed $\bar{v}$; for copper $\bar{v} = 1.6 \times 10^8$ cm/sec. In the absence of an electron field, the directions in which the electrons move are completely random, like those of the molecules of a gas confined to a container.

The electrons collide constantly with the ionic cores of the conductor, that is, they interact with the lattice, often suffering sudden changes in speed and direction. These collisions remind us of the collisions of gas molecules confined to a container. As in the case of molecular collisions, we can describe electron-lattice collisions by a *mean free path* λ, where λ is the average distance that an electron travels between collisions.*

In an ideal metallic crystal at 0°K electron-lattice collisions would not occur, according to the predictions of quantum physics, that is, $\lambda \to \infty$ as $T \to 0$°K for ideal crystals. Collisions take place in actual crystals because (*a*) the ionic cores at any temperature T are vibrating about their equilibrium positions in a random way, (*b*) impurities, that is, foreign atoms, may be present, and (*c*) the crystal may contain lattice imperfections, such as rows of missing atoms and displaced atoms. On this view it is not surprising that the resistivity of a metal can be increased by (*a*) raising its temperature, (*b*) adding small amounts of impurities, and (*c*) straining it severely, as by drawing it through a die, to increase the number of lattice imperfections.

When an electric field is applied to a metal, the electrons modify their random motion in such a way that they drift slowly, in the opposite direction to that of the field, with an average drift speed v_d. This drift speed is much less than the effective average speed $\bar{v}$ mentioned above (see Example 2). Figure 31-7 suggests the relationship between these two speeds. The solid lines suggest a possible random path followed by an electron in the absence of an applied field; the electron proceeds from x to y, making six collisions on the way. The dashed curves show how this same event *might* have occurred if an electric field **E** had been applied. Note that the electron drifts steadily to the right, ending at y' rather than at y. In preparing Fig. 31-7, it has been assumed that the drift speed v_d is $0.02\bar{v}$; actually, it is more like $10^{-10}\bar{v}$, so that the "drift" exhibited in the figure is greatly exaggerated.

The drift speed v_d can be calculated in terms of the applied electric field E and of $\bar{v}$ and λ. When a field is applied to an electron in the metal it will experience a force eE which will impart to it an acceleration a given by Newton's second law,

$$a = \frac{eE}{m}.$$

---

* It can be shown that collisions between electrons occur only rarely and have little effect on the resistivity.

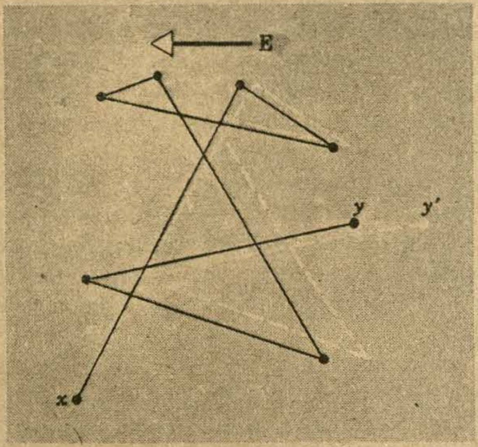


Fig. 31–7 The solid lines show an electron moving from x to y, making six collisions. The dashed curves show what the electron path *might* have been in the presence of an electric field **E**. Note the steady drift in the direction of $-\mathbf{E}$.

Consider an electron that has just collided with an ion core. The collision, in general, will momentarily destroy the tendency to drift and the electron will have a truly random direction after the collision. At its next collision the electron's velocity will have changed, on the average, by $a(l/\bar{v})$ where $l/\bar{v}$ is the mean time between collisions. We call this the drift speed v_d, or

$$v_d = a\left(\frac{\lambda}{\bar{v}}\right) = \frac{eE\lambda}{m\bar{v}}. \tag{31–13}$$

The electron's motion through the conductor is analogous to the constant rate of fall of a stone in water. The gravitational force F_g on the stone is opposed by a viscous resisting force that is proportional to the velocity, or

$$F_g = mg = bv,$$

where b is a viscous coefficient (see Section 15–8). Thus the constant terminal speed of the stone is

$$v = \left(\frac{1}{b}\right)F_g.$$

We can rewrite Eq. 31–13 as

$$v_d = \left(\frac{\lambda}{m\bar{v}}\right)F_E,$$

where F_E ($= eE$) is the electrical force. Comparison of these equations shows that the equivalent "viscous coefficient" for the motion of an electron in a particular conductor is $\lambda/2m\bar{v}$. If λ is short, the conductor exhibits a greater "viscous effect" on the electron motion, and the drift speed v_d is proportionally lower.

We may express v_d in terms of the current density (Eq. 31–5) and combine with Eq. 31–13 to obtain

$$v_d = \frac{j}{ne} = \frac{eE\lambda}{m\bar{v}}.$$

Combining this with Eq. 31–7 ($\rho = E/j$) leads finally to

$$\rho = \frac{m\bar{v}}{ne^2\lambda}. \tag{31–14}$$

Equation 31–14 can be taken as a statement that metals obey Ohm's law *if* we can show that $\bar{v}$ and λ do not depend on the applied electric field E. In this case ρ will not depend on E, which (see Section 31–3) is the criterion that a material obey Ohm's law. The quantities $\bar{v}$ and λ depend on the speed distribution of the conduction electrons. We have seen that this distribution is affected only slightly by the application of even a relatively large electric field, since $\bar{v}$ is of the order of 10^8 cm/sec and v_d (see Example 1) only of the order of 10^{-2} cm/sec, a ratio of 10^{10}. We may be sure that whatever the values of $\bar{v}$ and λ are (for copper at 20°C, say) in the absence of a field they remain essentially unchanged when the field is applied. Thus the right side of Eq. 31–14 is independent of E and the material obeys Ohm's law. The numerical calculation of ρ from Eq. 31–14 is hampered by the difficulty of calculating λ, although the calculation has been carried out in a number of cases.

▶ **Example 4.** What are (*a*) the mean time τ between collisions and (*b*) the mean free path for free electrons in copper?

(*a*) From Eq. 31–14 (see also Example 2), we have

$$\tau = \frac{\lambda}{\bar{v}} = \frac{m}{ne^2\rho} = \frac{(9.1 \times 10^{-31}\text{ kg})}{(8.4 \times 10^{28}/\text{meter}^3)(1.6 \times 10^{-19}\text{ coul})^2(1.7 \times 10^{-8}\text{ ohm-m})}$$
$$= 2.5 \times 10^{-14}\text{ sec}.$$

(*b*) The mean free path is

$$\lambda = \tau\bar{v} = (2.5 \times 10^{-14}\text{ sec})(1.6 \times 10^8\text{ cm/sec}) = 4.0 \times 10^{-6}\text{ cm}.$$

This is about 200 ionic diameters. ◀

31–5 Energy Transfers in an Electric Circuit

Figure 31–8 shows a circuit consisting of a battery B connected to a "black box." A steady current i exists in the connecting wires and a steady potential difference V_{ab} exists between the terminals a and b. The box might contain a resistor, a motor, or a storage battery, among other things.

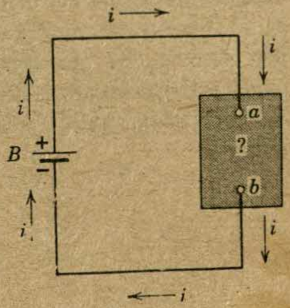


Fig. 31–8 A battery B sets up a current in a circuit containing a "black box."

Terminal a, connected to the positive battery terminal, is at a higher potential than terminal b. If a charge dq moves through the box from a to b, this charge will decrease its electric potential energy by $dq\ V_{ab}$ (see Section 29–6). The conservation-of-energy principle tells us that this energy is transferred in the box from electric potential energy to some other form. What that other form will be depends on what is in the box. In a time dt the energy dU transferred inside the box is then

$$dU = dq\ V_{ab} = i\ dt\ V_{ab}.$$

We find the *rate* of energy transfer P by dividing by the time, or

$$P = \frac{dU}{dt} = iV_{ab}. \tag{31-15}$$

If the device in the box is a motor, the energy appears largely as mechanical work done by the motor; if the device is a storage battery that is being charged, the energy appears largely as stored chemical energy in this second battery.

If the device is a resistor, we assert that the energy appears as heat in the resistor. To see this, consider a stone of mass m that falls through a height h. It decreases its gravitational potential energy by mgh. If the stone falls in a vacuum or—for practical purposes—in air, this energy is transformed into kinetic energy of the stone. If the stone falls in water, however, its speed eventually becomes constant, which means that the kinetic energy no longer increases. The potential energy that is steadily being made available as the stone falls then appears as thermal energy in the stone and the surrounding water. It is the viscous, friction-like drag of the water on the surface of the stone that stops the stone from accelerating, and it is at this surface that thermal energy appears.

The course of the electrons through the resistor is much like that of the stone through water. The electrons travel with a constant drift speed v_d and thus do not gain kinetic energy. The electric potential energy that they lose is transferred to the resistor as heat. On a microscopic scale this can be understood in that collisions between the electrons and the lattice (see Fig. 21–5) increase the amplitude of the thermal vibrations of the lattice; on a macroscopic scale this corresponds to a temperature increase. This effect, which is thermodynamically irreversible, is called *Joule heating*.

For a resistor we can combine Eqs. 31–15 and 31–6 ($R = V/i$) and obtain either

$$P = i^2R \tag{31-16}$$

or

$$P = \frac{V^2}{R}. \tag{31-17}$$

Note that Eq. 31–15 applies to electrical energy transfer of *all* kinds; Eqs. 31–16 and 31–17 apply only to the transfer of electrical energy to heat energy in a resistor. Equations 31–16 and 31–17 are known as *Joule's law*. This law is a particular way of writing the conservation-of-energy principle for the special case in which electrical energy is transferred into heat energy.

The unit of power that follows from Eq. 31–15 is the volt-amp. It can be written as

$$\begin{aligned}1\text{ volt-amp} &= 1\ \cancel{\text{volt}}\text{-}\cancel{\text{amp}}\left(\frac{1\text{ joule}}{1\ \cancel{\text{volt}}\times 1\ \cancel{\text{coul}}}\right)\left(\frac{1\ \cancel{\text{coul}}}{1\ \cancel{\text{amp}}\times 1\text{ sec}}\right)\\ &= 1\text{ joule/sec.}\end{aligned}$$

The first conversion factor in parenthesis comes from the definition of the volt (Eq. 29–1); the second comes from the definition of the coulomb. The joule/sec is such a common unit that it is given a special name of its own, the *watt;* see Section 7–7. Power is not an exclusively electrical concept, of course, and we can express in watts the power ($= \mathbf{F}\cdot\mathbf{v}$) expended by an agent that exerts a force **F** while it moves with a velocity **v**.

► **Example 5.** You are given a 20-ft length of heating wire made of the special alloy Nichrome; it has a resistance of 24 ohms. Can you obtain more heat by winding one coil or by cutting the wire in two and winding two separate coils? In each case the coils are to be connected individually across a 110-volt line.

The power P for the single coil is given by Eq. 31–17:

$$P = \frac{V^2}{R} = \frac{(110\text{ volts})^2}{24\text{ ohms}} = 500\text{ watts.}$$

The power for a coil of half the length is given by

$$P' = \frac{(110\text{ volts})^2}{12\text{ ohms}} = 1000\text{ watts.}$$

There are two "half-coils," so that the total power obtained by cutting the wire in half is 2000 watts, or four times that for the single coil. This would seem to suggest that we could buy a 500-watt heating coil, cut it in half, and rewind it to obtain 2000 watts. Why is this not a practical idea? ◄

QUESTIONS

1. Name other physical quantities that, like current, are scalars having a sense represented by an arrow in a diagram.

2. What conclusions can you draw by applying Eq. 31–4 to a closed surface through which a number of wires pass in random directions, carrying steady currents of different sizes?

3. A potential difference V is applied to a circular cylinder of carbon by clamping it between circular copper electrodes, as in Fig. 31–9. Discuss the difficulty of calculating the resistance of the carbon cylinder, using the relation $R = \rho L/A$.

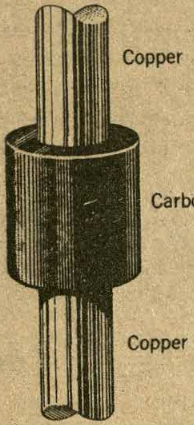


Fig. 31–9

4. How would you measure the resistance of a pretzel-shaped conductor? Give specific details to clarify the concept.

5. Discuss the difficulties of testing whether the filament of a light bulb obeys Ohm's law.

6. Does the relation $V = iR$ apply to nonohmic resistors?

7. The temperature coefficient of resistance of a thermistor is negative and varies greatly with temperature. Account qualitatively for the shape of the curve of i versus V for the thermistor of Fig. 31–6.

8. A potential difference V is applied to a copper wire of diameter d and length l. What is the effect on the electron drift speed of (*a*) doubling V, (*b*) doubling l, and (*c*) doubling d?

9. If the drift speeds of the electrons in a conductor under ordinary circumstances are so slow (see Example 2), why do the lights in a room turn on so quickly after the switch is closed?

10. Can you think of a way to measure the drift speed for electrons by timing their travel along a conductor?

11. Why are the dashed white lines in Fig. 31–7 curved slightly?

12. A current i enters the top of a copper sphere of radius R and leaves at a diametrically opposite point. Are all parts of the sphere equally effective in dissipating Joule heat?

13. What special characteristics must (*a*) heating wire and (*b*) fuse wire have?

14. Equation 31-16 ($P = i^2R$) seems to suggest that the rate of Joule heating in a resistor is reduced if the resistance is made less; Eq. 31-17 ($P = V^2/R$) seems to suggest just the opposite. How do you reconcile this apparent paradox?

15. Is the filament resistance lower or higher in a 500-watt light bulb than in a 100-watt bulb? Both bulbs are designed to operate on 110 volts.

16. Five wires of the same length and diameter are connected in turn between two points maintained at constant potential difference. Will heat be developed at the fastest rate in the wire of (a) the smallest or (b) the largest resistance?

PROBLEMS

1. A current of 5 amp exists in a 10-ohm resistance for 4 min. (a) How many coulombs and (b) how many electrons pass through any cross section of the resistor in this time?

2. A current is established in a gas discharge tube when a sufficiently high potential difference is applied across the two electrodes in the tube. The gas ionizes; electrons move toward the positive terminal and positive ions toward the negative terminal. What are the magnitude and sense of the current in a hydrogen discharge tube in which 3.1×10^{18} electrons and 1.1×10^{18} protons move past a cross-sectional area of the tube each second?

3. A copper wire and an iron wire of the same length have the same potential difference applied to them. (a) What must be the ratio of their radii if the current is to be the same? (b) Can the current density be made the same by suitable choices of the radii?

4. A current i enters one corner of a square sheet of copper and leaves at the opposite corner. Sketch arrows for various points within the square to represent the relative values of $\mathbf{j}$. Intuitive guesses rather than detailed mathematical analysis are called for.

5. The belt of an electrostatic generator is 50 cm wide and travels at 30 meters/sec. The belt carries charge into the sphere at a rate corresponding to 10^{-4} amp. Compute the surface charge density on the belt.

6. A square aluminum rod is 1.0 meter long and 5.0 mm on edge. (a) What is the resistance between its ends? (b) What must be the diameter of a circular 1.0-meter copper rod if its resistance is to be the same?

7. A wire with a resistance of 6.0 ohms is drawn out so that its new length is three times its original length. Find the resistance of the longer wire, assuming that the resistivity and density of the material are not changed during the drawing process.

8. A copper wire and an iron wire of equal length l and diameter d are joined and a potential difference V is applied between the ends of the composite wire. Calculate (a) the electric field strength in each wire, (b) the current density in each wire, and (c) the potential difference across each wire. Assume that $l = 10$ meters, $d = 2.0$ mm, and $V = 100$ volts.

9. A rod of a certain metal is 1.00 meter long and 0.550 cm in diameter. The resistance between its ends (at 20°C) is 2.87×10^{-3} ohm. A round disk is formed of this same material, 2.00 cm in diameter and 1.00 mm thick. (a) What is the resistance between the opposing round faces? (b) What is the material?

10. Steel trolley-car rail has a cross-sectional area of 7.1 in.^{2} What is the resistance of 10 miles of single track? The resistivity of the steel is 6.0×10^{-7} ohm-m.

11. (a) At what temperature would the resistance of a copper conductor be double its resistance at 0°C? (b) Does this same temperature hold for all copper conductors, regardless of shape or size?

12. It is desired to make a long cylindrical conductor whose temperature coefficient of resistivity at 20°C will be close to zero. (a) If such a conductor is made by assembling alternate disks of iron and carbon, what is the ratio of the thickness of a carbon disk to that of an iron disk? Assume that the temperature remains essentially the same in each disk. (b) What is the ratio of the rate of Joule heating in a carbon disk to that in an iron disk?

13. When a metal rod is heated, not only its resistance but also its length and its cross-sectional area change. The relation $R = \rho l/A$ suggests that all three factors should be taken into account in measuring ρ at various temperatures. If the temperature changes by 1.0 C°, what per cent changes in R, l, and A occur for a copper conductor. What conclusion do you draw? The coefficient of linear expansion is 1.7×10^{-5}/C°.

14. The copper windings of a motor have a resistance of 50 ohms at 20°C, when the motor is idle. After running for several hours the resistance rises to 58 ohms. What is the temperature of the windings?

15. (*a*) Using data from Fig. 31–5, plot the resistance of the vacuum tube as a function of applied potential difference. (*b*) Repeat for the thermistor of Fig. 31–6.

16. A small but measurable current of 10^{-10} amp exists in a copper wire whose diameter is 0.10 in. Calculate the electron drift speed.

17. Heat is developed in a resistor at a rate of 100 watts when the current is 3.0 amp. What is the resistance in ohms?

18. A potential difference of 1.0 volt is applied to a 100-ft length of #18 copper wire (diameter = 0.040 in.). Calculate (*a*) the current, (*b*) the current density, (*c*) the electric field strength, and (*d*) the rate of Joule heating.

19. The National Board of Fire Underwriters has fixed safe current-carrying capacities for various sizes and types of wire. For #10 rubber-coated copper wire (wire diameter = 0.10 in.) the maximum safe current is 25 amp. At this current, find (*a*) the current density, (*b*) the electric field strength, (*c*) the potential difference for 1000 ft of wire, and (*d*) the rate of Joule heating for 1000 ft of wire.

20. A 500-watt immersion heater is placed in a pot containing 2.0 liters of water at 20°C. (*a*) How long will it take to bring the water to boiling temperature, assuming that 80% of the available energy is absorbed by the water? (*b*) How much longer will it take to boil half the water away?

21. A nichrome heater dissipates 500 watts when the applied potential difference is 110 volts and the wire temperature is 800°C. How much power would it dissipate if the wire temperature were held to 200°C by immersion in a bath of cooling oil? The applied potential difference remains the same; $\bar{\alpha}$ for nichrome is about 4×10^{-4}/C°.

22. A beam of 16-Mev deuterons from a cyclotron falls on a copper block. The beam is equivalent to a current of 15×10^{-6} amp. (*a*) At what rate do deuterons strike the block? (*b*) At what rate is heat produced in the block?

23. A "500-watt" heating unit is designed to operate from a 115-volt line. (*a*) By what percentage will its heat output drop if the line voltage drops to 110 volts? Assume no change in resistance. (*b*) Taking the variation of resistance with temperature into account, would the actual heat output drop be larger or smaller than that calculated in (*a*)?

24. Show that p, the power per unit volume transformed into Joule heat in a resistor, can be written as

$$p = j^2\rho \qquad \text{or} \qquad p = E^2/\rho.$$

Electromotive Force and Circuits

CHAPTER 32

32–1 Electromotive Force

There exist in nature certain devices such as batteries and electric generators which are able to maintain a potential difference between two points to which they are attached. Such devices are called seats of *electromotive force* (abbr. emf). In this chapter we do not discuss their internal construction or detailed mode of action but confine ourselves to describing their gross electrical characteristics and to exploring their usefulness in electric circuits.

Figure 32–1a shows a seat of emf B, represented by a battery, connected to a resistor R. The seat of emf maintains its upper terminal positive and its lower terminal negative, as shown by the + and − signs. In the circuit external to B positive charge carriers would be driven in the direction shown by the arrows marked i. In other words, a clockwise current would be set up.

An emf is represented by an arrow which is placed next to the seat and points in the direction in which the seat, acting alone, would cause a positive charge carrier to move in the external circuit. A small circle is drawn on the tail of an emf arrow so that it will not be confused with a current arrow.

A seat of emf must be able to do work on charge carriers that enter it. In the circuit of Fig. 32–1a, for example, the seat acts to move positive charges from a point of low potential (the negative terminal) through the seat to a point of high potential (the positive terminal). This reminds us of a pump, which can cause water to move from a place of low gravitational potential to a place of high potential.

In Fig. 32–1a a charge dq passes through *any* cross section of the circuit in time dt. In particular, this charge enters the seat of emf $\mathcal{E}$ at its low-potential

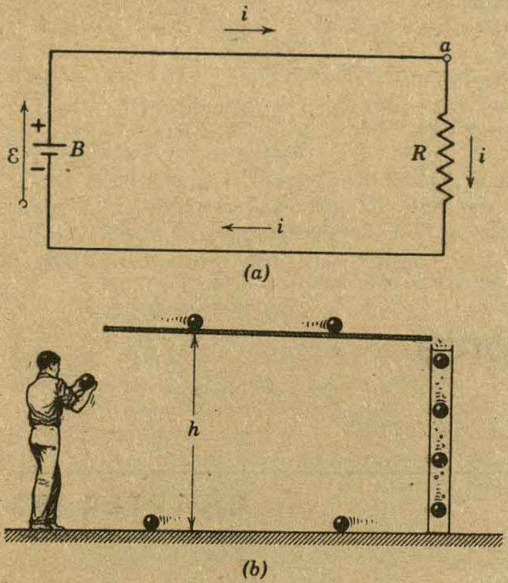


Fig. 32-1 (*a*) A simple electric circuit and (*b*) its gravitational analog.

end and leaves at its high-potential end. The seat must do an amount of work dW on the (positive) charge carriers to force them to go to the point of higher potential. The emf $\mathcal{E}$ of the seat is defined from

$$\mathcal{E} = dW/dq. \tag{32-1}$$

The unit of emf is the joule/coul (see Eq. 29–1) which is the *volt*. We might be inclined to say that a battery has an emf of 1 volt if it maintains a difference of potential of 1 volt between its terminals. This is true only under certain conditions, which we describe in Section 32–4.

If a seat of emf does work on a charge carrier, energy must be transferred within the seat. In a battery, for example, chemical energy is transferred into electrical energy. Thus we can describe a seat of emf as a device in which chemical, mechanical, or some other form of energy is changed (reversibly) into electrical energy. The chemical energy provided by the battery in Fig. 32–1*a* is stored in the electric and the magnetic [*] fields that surround the circuit. This stored energy does not increase because it is being drained away, by transfer to Joule heat in the resistor, at the same rate at which it is supplied. The electric and magnetic fields play an intermediary role in the energy transfer process, acting as a storage reservoir.

Figure 32–1*b* shows a gravitational analog of Fig. 32–1*a*. In the top figure the seat of emf B does work on the charge carriers. This energy, stored temporarily as electromagnetic field energy, appears eventually as Joule heat in resistor R. In the lower figure the man, in lifting the bowling balls from

---

[*] A current in a wire is surrounded by a magnetic field, and this field, like the electric field, can also be viewed as a site of stored energy (see Section 36–4).

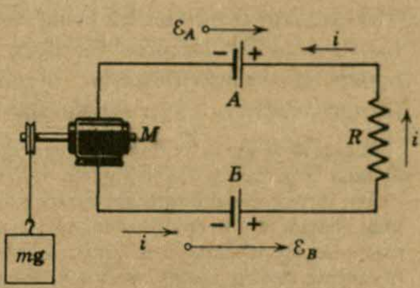
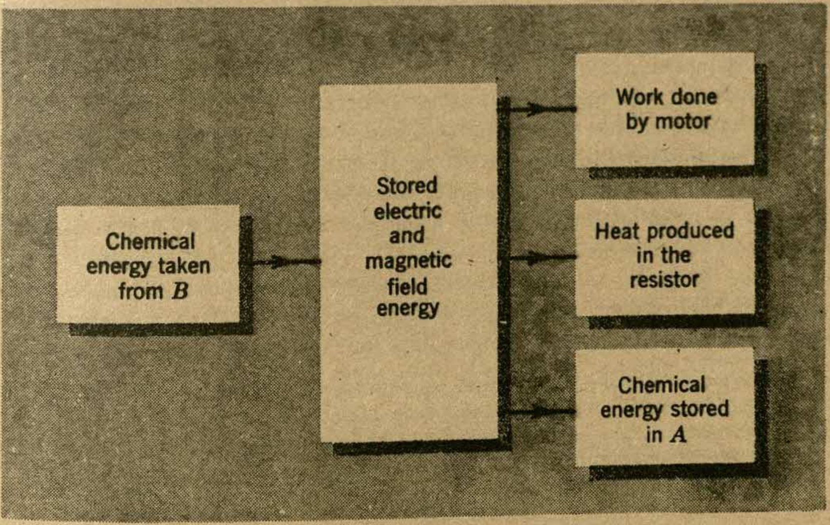


Fig. 32-2 Two batteries, a resistor, and a motor, connected in a single-loop circuit. It is given that $\mathcal{E}_B > \mathcal{E}_A$.

the floor to the shelf, does work on them. This energy is stored temporarily as gravitational field energy. The balls roll slowly and uniformly along the shelf, dropping from the right end into a cylinder full of viscous oil. They sink to the bottom at constant speed, are removed by a trapdoor mechanism not shown, and roll back along the floor to the left. The energy put into the system by the man appears eventually as heat in the viscous fluid. The energy supplied by the man comes from his own internal (chemical) energy. The circulation of charges in Fig. 32-1*a* will stop eventually if battery *B* is not charged; the circulation of bowling balls in Fig. 32-1*b* will stop eventually if the man does not replenish his store of internal energy by eating.

Figure 32-2 shows a circuit containing two (ideal) batteries, *A* and *B*, a resistor *R*, and an (ideal) electric motor employed in lifting a weight. The batteries are connected so that they tend to send charges around the circuit in opposite directions; the actual direction of the current is determined by *B*, which supplies the larger potential difference. The energy transfers in this circuit are


Chemical energy taken from *B*
Stored electric and magnetic field energy
Work done by motor
Heat produced in the resistor
Chemical energy stored in *A*

The chemical energy in B is being steadily depleted, the energy appearing in the three forms shown on the right. Battery A is being "charged" while battery B is being discharged. Again, the electric and magnetic fields that surround the circuit act as an intermediary.

It is part of the definition of an emf that the energy transfer process be *reversible*, at least in principle. The student will recall that a reversible process is one that passes through equilibrium states; its course can be reversed by making an infinitesimal change in the environment of the system; see Section 25–2. A battery, for example, can either be on charge or discharge; a generator can be driven mechanically, producing electric energy, or it can be operated backward as a motor. The (reversible) energy transfers here are

$$\text{electrical} \rightleftharpoons \text{chemical}$$

and

$$\text{electrical} \rightleftharpoons \text{mechanical}.$$

Joule heating is an electric energy transfer that is *not* reversible. We can easily heat a conductor by supplying electric energy to it, but it is not possible to set up a current in a closed copper loop by heating the loop uniformly. Because of this lack of reversibility, we do not associate an emf with Joule heating.

32–2 Calculating the Current

In a time dt an amount of energy given by $i^2R\,dt$ will appear in the resistor of Fig. 32–1a as Joule heat. During this same time a charge dq ($= i\,dt$) will have moved through the seat of emf, and the seat will have done work on this charge (see Eq. 32–1) given by

$$dW = \mathcal{E}dq = \mathcal{E}i\,dt.$$

From the conservation of energy principle, the work done by the seat must equal the Joule heat, or

$$\mathcal{E}i\,dt = i^2R\,dt.$$

Solving for i, we obtain

$$i = \mathcal{E}/R. \tag{32–2}$$

We can also derive Eq. 32–2 by considering that if electric potential is to have any meaning a given point can have only one value of potential at any given time. If we start at any point in the circuit of Fig. 32–1a and, in imagination, go around the circuit in either direction, adding up algebraically the changes in potential that we encounter, we must arrive at the same potential when we return to our starting point. In other words, *the algebraic sum of the changes in potential encountered in a complete traversal of the circuit must be zero.*

In Fig. 32–1a let us start at point a, whose potential is V_a,* and traverse the circuit clockwise. In going through the resistor, there is a change in potential of $-iR$. The minus sign shows that the top of the resistor is higher in potential than the bottom, which must be true, because positive charge carriers move of their own accord from high to low potential. As we traverse the battery from bottom to top, there is an *increase* of potential $+\mathcal{E}$ because the battery does (positive) work on the charge carriers, that is, it moves them from a point of low potential to one of high potential. Adding the algebraic sum of the changes in potential to the initial potential V_a must yield the identical value V_a, or

---

* The actual value of V_a depends on assumptions made in the definition of potential (as described in Section 29–1). The numerical value of V_a is not important because, as in most electric circuit situations, we are concerned here with *differences* of potential. Point a in Fig. 32–1a (or any other single point in that figure) could be connected to ground (symbol ⏚) and assigned the potential $V_a = 0$, following a common practice.

$$V_a - iR + \mathcal{E} = V_a.$$

We write this as

$$-iR + \mathcal{E} = 0,$$

which is independent of the value of V_a and which asserts explicitly that the algebraic sum of the potential changes for a complete circuit traversal is zero. This relation leads directly to Eq. 32-2.

These two ways to find the current in single-loop circuits, based on the conservation of energy and on the concept of potential, are completely equivalent because potential differences are defined in terms of work and energy (Section 29-1). The statement that the sum of the changes in potential encountered in making a complete loop is zero is called *Kirchhoff's second rule;* for brevity we call it the *loop theorem.* It must always be borne in mind that this theorem is simply a particular way of stating the law of conservation of energy for electric circuits.

To prepare for the study of more complex circuits, let us examine the rules for finding potential differences; these rules follow from the previous discussion. They are not meant to be memorized but rather to be so thoroughly understood that it becomes trivial to re-derive them on each application.

1. If a resistor is traversed in the direction of the current, the change in potential is $-iR$; in the opposite direction it is $+iR$.
2. If a seat of emf is traversed in the direction of the emf, the change in potential is $+\mathcal{E}$; in the opposite direction it is $-\mathcal{E}$.

32-3 Other Single-Loop Circuits

Figure 32-3*a* shows a circuit which emphasizes that all seats of emf have an intrinsic internal resistance r. This resistance cannot be removed—although we would usually like to do so—because it is an inherent part of the device. The figure shows the internal resistance r and the emf separately, although, actually, they occupy the same region of space.

If we apply the loop theorem, starting at b and going around clockwise, we obtain

$$V_b + \mathcal{E} - ir - iR = V_b$$

or

$$+\mathcal{E} - ir - iR = 0.$$

The student should compare these equations with Fig. 32-2*b*, which shows the changes in potential graphically. In writing these equations, note that we traversed r and R in the direction of the current and $\mathcal{E}$ in the direction of the

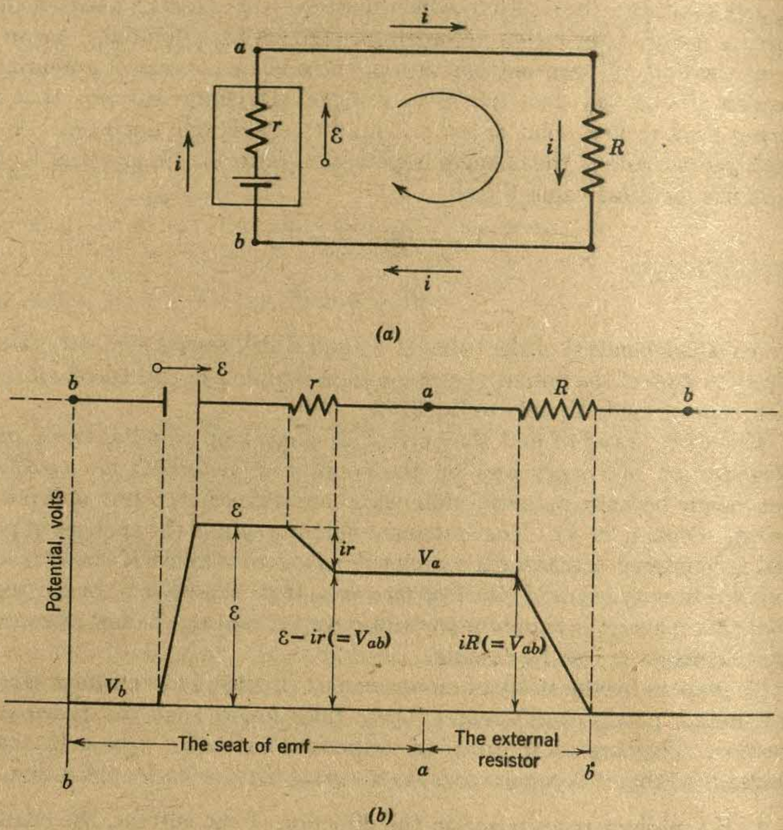


Fig. 32-3 A single-loop circuit. The rectangular block is a seat of emf with internal resistance r. (b) The same circuit is drawn for convenience as a straight line. Directly below are shown the changes in potential that one encounters in traversing the circuit clockwise from point b.

emf. The same equation follows if we start at any other point in the circuit or if we traverse the circuit in a counterclockwise direction. Solving for i gives

$$i = \frac{\mathcal{E}}{R + r}. \tag{32-3}$$

▶ **Example 1.** *Resistors in series.* Resistors in series are connected so that there is only one conducting path through them, as in Fig. 32-4. What is the equivalent resistance R of this series combination? The equivalent resistance is the single resistance R which, substituted for the series combination between the terminals ab, will leave the current i unchanged.

Applying the loop theorem (going clockwise from a) yields

$$-iR_1 - iR_2 - iR_3 + \mathcal{E} = 0$$

or

$$i = \frac{\mathcal{E}}{R_1 + R_2 + R_3}.$$

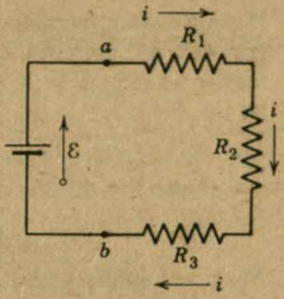


Fig. 32–4 Example 1. Three resistors are connected in series between terminals a and b.

For the equivalent resistance R

$$i = \frac{\mathcal{E}}{R}$$

or

$$R = R_1 + R_2 + R_3. \tag{32–4}$$

The extension to more than three resistors is clear. ◀

32–4 Potential Differences

We often want to compute the potential difference between two points in a circuit. In Fig. 32–3a for example, what is the relationship between the potential difference V_{ab} $(= V_a - V_b)$ between points b and a and the fixed circuit parameters $\mathcal{E}$, r, and R? To find this relationship, let us start from point b and traverse the circuit to point a, passing through resistor R against the current. If V_b and V_a are the potentials at b and a, respectively, we have

$$V_b + iR = V_a$$

because we experience an increase in potential in traversing a resistor against the current arrow. We rewrite this relation as

$$V_{ab} = V_a - V_b = +iR,$$

which tells us that V_{ab}, the potential difference between points a and b, has the magnitude iR and that point a is more positive than point b. Combining this last equation with Eq. 32–3 yields

$$V_{ab} = \mathcal{E}\frac{R}{R + r}. \tag{32–5}$$

To sum up: To find the potential difference between any two points in a circuit start at one point and traverse the circuit to the other, following any path, and add up algebraically the potential changes encountered. This algebraic sum will be the potential difference. This procedure is similar to that for finding the current in a closed loop, except that here the potential differences are added up over part of a loop and not over the whole loop.

The potential difference between any two points can have only one value; thus we must obtain the same answer for all paths that connect these points.

If we consider two points on the side of a hill, the measured difference in gravitational potential (that is, in altitude) between them is the same no matter what path is followed in going from one to the other. In Fig. 32–3*a* let us calculate V_{ab}, using a path passing through the seat of emf. We have

$$V_b + \mathcal{E} - ir = V_a$$

or (see also Fig. 32–3*b*)

$$V_{ab} = V_a - V_b = +\mathcal{E} - ir.$$

Again, combining with Eq. 32–3 leads to Eq. 32–5.

The terminal potential difference of the battery V_{ab}, as Eq. 32–5 shows, is less than $\mathcal{E}$ unless the battery has no internal resistance ($r = 0$) or if it is on open circuit ($R = \infty$); then V_{ab} is equal to $\mathcal{E}$. Thus the emf of a device is equal to its terminal potential difference *when on open circuit.*

▶ **Example 2.** In Fig. 32–5*a* let $\mathcal{E}_1$ and $\mathcal{E}_2$ be 2.0 volts and 4.0 volts, respectively; let the resistances r_1, r_2, and R be 1.0 ohm, 2.0 ohms, and 5.0 ohms, respectively. What is the current?

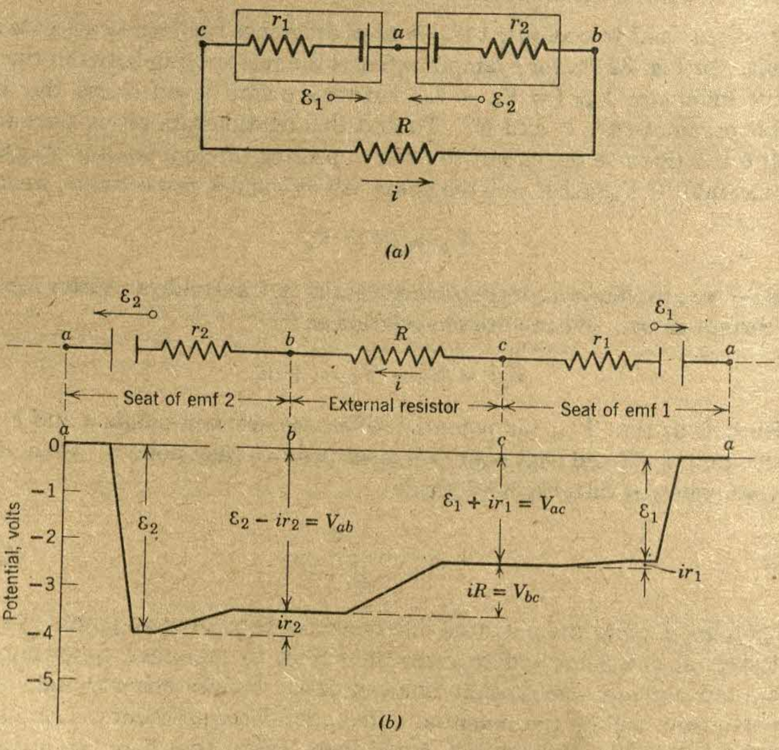


Fig. 32–5 (*a*) A single-loop circuit. (*b*) The same circuit is shown schematically as a straight line, the potential differences encountered in traversing the circuit clockwise from point *a* being displayed directly below. In the lower figure the potential of point *a* was assumed to be zero for convenience.

Emfs $\mathcal{E}_1$ and $\mathcal{E}_2$ oppose each other, but because $\mathcal{E}_2$ is larger it controls the direction of the current. Thus i will be counterclockwise. The loop theorem, going clockwise from a, yields

$$-\mathcal{E}_2 + ir_2 + iR + ir_1 + \mathcal{E}_1 = 0.$$

The student should check that the same result is obtained by going around counterclockwise. He should also compare this equation carefully with Fig. 32-5*b*, which shows the potential changes graphically.

Solving for i yields

$$i = \frac{\mathcal{E}_2 - \mathcal{E}_1}{R + r_1 + r_2} = \frac{4.0\text{ volts} - 2.0\text{ volts}}{5.0\text{ ohms} + 1.0\text{ ohm} + 2.0\text{ ohms}}$$

$$= 0.25\text{ amp}.$$

It is not necessary to know in advance what the actual direction of the current is. To show this, let us assume that the current in Fig. 32-5*a* is clockwise, an assumption that we know is incorrect. The loop theorem then yields (going clockwise from a)

$$-\mathcal{E}_2 - ir_2 - iR - ir_1 + \mathcal{E}_1 = 0$$

or

$$i = \frac{\mathcal{E}_1 - \mathcal{E}_2}{R + r_1 + r_2}.$$

Substituting numerical values (see above) yields -0.25 amp for the current. The minus sign tells us that the current is in the opposite direction from the one we have assumed.

In more complex circuit problems involving many loops and branches it is often impossible to know in advance the correct directions for the currents in all parts of the circuit. We can assume directions for the currents at random. Those currents for which positive numerical values are obtained will have the correct directions; those for which negative values are obtained will be exactly opposite to the assumed directions. In all cases the numerical values will be correct.

Example 3. What is the potential difference (*a*) between points b and a in Fig. 32-5*a*? (*b*) Between points a and c?

(*a*) For points a and b we start at b and traverse the circuit to a, obtaining

$$V_{ab}\,(= V_a - V_b) = -ir_2 + \mathcal{E}_2 = -(0.25\text{ amp})(2.0\text{ ohms}) + 4.0\text{ volts}$$

$$= +3.5\text{ volts}.$$

Thus a is more positive than b and the potential difference (3.5 volts) is *less than* the emf (4.0 volts); see Fig. 32-5*b*.

(*b*) For points c and a, we start at c and traverse the circuit to a, obtaining

$$V_{ac}\,(= V_a - V_c) = +\mathcal{E}_1 + ir_1 = +2.0\text{ volts} + (0.25\text{ amp})(1.0\text{ ohm})$$

$$= +2.25\text{ volts}.$$

This tells us that a is at a higher potential than c. The terminal potential difference of $\mathcal{E}_1$ (2.25 volts) is *larger than* the emf (2.0 volts); see Fig. 32-5*b*. Charge is being forced through $\mathcal{E}_1$ in a direction opposite to the one in which it would send charge if it were acting by itself; if $\mathcal{E}_1$ is a storage battery, it is being charged at the expense of $\mathcal{E}_2$.

Let us test the first result by proceeding from b to a along a different path, namely, through R, r_1, and $\mathcal{E}_1$. We have

$$V_{ab} = iR + ir_1 + \mathcal{E}_1 = (0.25\text{ amp})(5.0\text{ ohms}) + (0.25\text{ amp})(1.0\text{ ohm}) + 2.0\text{ volts} = +3.5\text{ volts},$$

which is the same as the earlier result. ◀

32-5 Multiloop Circuits

Figure 32–6 shows a circuit containing two loops. For simplicity, we have neglected the internal resistances of the batteries. There are two *junctions*, b and d, and three *branches* connecting these junctions. The branches are the left branch bad, the right branch bcd, and the central branch bd. If the emfs and the resistances are given, what are the currents in the various branches?

We label the currents in the branches as i_1, i_2, and i_3, as shown. Current i_1 has the same value for any cross section of the left branch from b to d. Similarly, i_2 has the same value everywhere in the right branch and i_3 in the central branch. The directions of the currents have been chosen arbitrarily. The careful reader will note that i_3 must point in a direction opposite to the one we have shown. We have deliberately drawn it in wrong to show how the formal mathematical procedures will always indicate this to us.

The three currents i_1, i_2, and i_3 carry charge either toward junction d or away from it. Charge does not accumulate at junction d, nor does it drain away from this junction because the circuit is in a steady-state condition. Thus charge must be removed from the junction by the currents at the same rate that it is brought into it. If we arbitrarily call a current approaching the junction positive and the one leaving the junction negative, then

$$i_1 + i_3 - i_2 = 0.$$

This equation suggests a general principle for the solution of multiloop circuits: *At any junction the algebraic sum of the currents must be zero.* This *junction theorem* is also known as *Kirchhoff's first rule.* Note that it is simply a statement of the conservation of charge. Thus our basic tools for solving circuits are (a) the conservation of energy (see p. 793) and (b) the conservation of charge.

For the circuit of Fig. 32–6, the junction theorem yields only one relationship among the three unknowns. Applying the theorem at junction b leads to exactly the same equation, as the student should verify. To solve for the three unknowns, we need two more independent equations; they can be found from the loop theorem.

In single-loop circuits there is only one conducting loop around which to apply the loop theorem, and the current is the same in all parts of this loop.

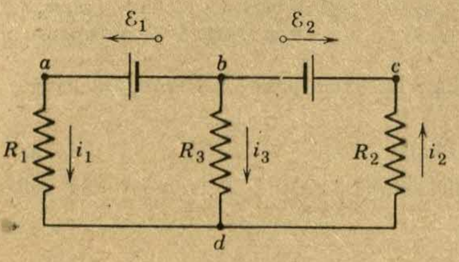


Fig. 32–6 A multiloop circuit.

In multiloop circuits there is more than one loop, and the current in general will not be the same in all parts of any given loop.

If we traverse the left loop of Fig. 32-6 in a counterclockwise direction, the loop theorem gives

$$\mathcal{E}_1 - i_1R_1 + i_3R_3 = 0. \tag{32-6}$$

The right loop gives

$$-i_3R_3 - i_2R_2 - \mathcal{E}_2 = 0. \tag{32-7}$$

These two equations, together with the relation derived earlier with the junction theorem, are the three simultaneous equations needed to solve for the unknowns i_1, i_2, and i_3. Doing so yields

$$i_1 = \frac{\mathcal{E}_1(R_2 + R_3) - \mathcal{E}_2R_3}{R_1R_2 + R_2R_3 + R_1R_3}, \tag{32-8a}$$

$$i_2 = \frac{\mathcal{E}_1R_3 - \mathcal{E}_2(R_1 + R_3)}{R_1R_2 + R_2R_3 + R_1R_3}, \tag{32-8b}$$

and

$$i_3 = \frac{-\mathcal{E}_1R_2 - \mathcal{E}_2R_1}{R_1R_2 + R_2R_3 + R_1R_3}. \tag{32-8c}$$

The student should supply the missing steps. Equation 32-8*c* shows that no matter what numerical values are given to the emfs and to the resistances the current i_3 will always have a negative value. This means that it will always point up in Fig. 32-6 rather than down, as we deliberately assumed. The currents i_1 and i_2 may be in either direction, depending on the particular numerical values given.

The student should verify that Eqs. 32-8 reduce to sensible conclusions in special cases. For $R_3 = \infty$, for example, we find

$$i_1 = i_2 = \frac{\mathcal{E}_1 - \mathcal{E}_2}{R_1 + R_2} \quad \text{and} \quad i_3 = 0.$$

What do these equations reduce to for $R_2 = \infty$?

The loop theorem can be applied to a large loop consisting of the entire circuit *abcda* of Fig. 32-6. This fact might suggest that there are more equations than we need, for there are only three unknowns and we already have three equations written in terms of them. However, the loop theorem yields for this loop

$$-i_1R_1 - i_2R_2 - \mathcal{E}_2 + \mathcal{E}_1 = 0,$$

which is nothing more than the sum of Eqs. 32-6 and 32-7. Thus this large loop does not yield another *independent* equation. It will never be found in solving multiloop circuits that there are more independent equations than variables.

▶ **Example 4.** *Resistors in parallel.* Figure 32–7 shows three resistors connected across the same seat of emf. Resistances across which the identical potential difference is applied are said to be in parallel. What is the equivalent resistance R of this parallel combination? The equivalent resistance is that single resistance which, substituted for the parallel combination between terminals ab, would leave the current i unchanged.

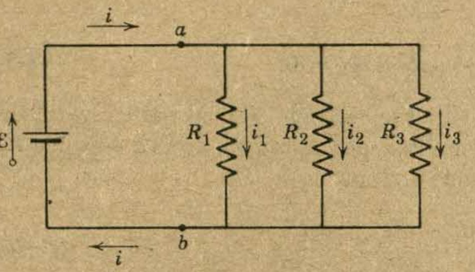


Fig. 32–7 Example 4. Three resistors are connected in parallel betweeen terminals a and b.

The currents in the three branches are

$$i_1 = \frac{V}{R_1}, \qquad i_2 = \frac{V}{R_2}, \qquad \text{and} \qquad i_3 = \frac{V}{R_3},$$

where V is the potential difference that appears between points a and b. The total current i is found by applying the junction theorem to junction a, or

$$i = i_1 + i_2 + i_3 = V\left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}\right).$$

If the equivalent resistance is used instead of the parallel combination, we have

$$i = \frac{V}{R}.$$

Combining these two equations gives

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}. \tag{32–9}$$

This formula can easily be extended to more than three resistances. Note that the equivalent resistance of a parallel combination is less than any of the resistances that make it up. ◀

32–6 Measuring Currents and Potential Differences

A meter to measure currents is called an *ammeter* (or a *milliammeter* or *microammeter*, depending on the size of the current to be measured). To determine the current in a wire, it is necessary to break or cut the wire and to insert the ammeter, so that the current to be measured passes through the meter (see Fig. 32–8).*

It is essential that the resistance R_A of the ammeter be *small* compared to other resistances in the circuit. Otherwise the act of inserting the meter will in itself change

---

* The meter must be connected so that the direction of current through it (assuming positive charge carriers) is *into* the meter terminal marked +. Otherwise the meter will deflect in a direction opposite to that intended.

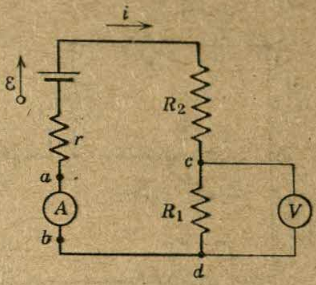


Fig. 32-8 An ammeter (A) is connected to read the current in a circuit, and a voltmeter (V) is connected to read the potential difference across resistor R_1.

the current to be measured. An ideal ammeter would have zero resistance. In the circuit of Fig. 32-8 the required condition, assuming that the voltmeter is not connected, is

$$R_A \ll r + R_1 + R_2.$$

A meter to measure potential differences is called a *voltmeter* (or a *millivoltmeter* or *microvoltmeter*). To find the potential difference between two points in a circuit, it is necessary to connect one of the voltmeter terminals to each of the circuit points, without breaking the circuit (see Fig. 32-8).*

It is essential that the resistance of the voltmeter R_V be *large* compared to any circuit resistance across which the voltmeter is connected. Otherwise the meter will itself constitute an important circuit element and will alter the circuit current and the potential difference to be measured. An ideal voltmeter would have an infinite resistance. In Fig. 32-8 the required condition is

$$R_V \gg R_1$$

In measuring potential difference in electronic circuits, where the effective circuit resistance may be of the order of 10^6 ohms or higher, it becomes necessary to use a *vacuum-tube voltmeter*, which is an electron-tube device designed specifically to have an extremely high effective resistance between its input terminals.

32-7 The Potentiometer

Figure 32-9 shows the rudiments of a *potentiometer*, which is a device for measuring an unknown emf $\mathcal{E}_x$. The currents and emfs are marked as shown. Applying the loop theorem to loop *abcd* yields

$$-\mathcal{E}_x - ir + (i_0 - i)R = 0,$$

where $i_0 - i$, by application of the junction theorem at a, is the current in resistor R. Solving for i yields

$$i = \frac{i_0R - \mathcal{E}_x,}{R + r}$$

in which R is a variable resistor. This relation shows that if R is adjusted to have the value R_x where

$$i_0R_x = \mathcal{E}_x, \qquad (32\text{-}10)$$

the current i in the branch *abcd* becomes zero. To *balance* the potentiometer in this way, R must be adjusted manually until the sensitive meter G reads zero.

---

* The voltmeter terminal marked + must be connected to the point of higher potential. Otherwise the meter will deflect in a direction opposite to that intended.

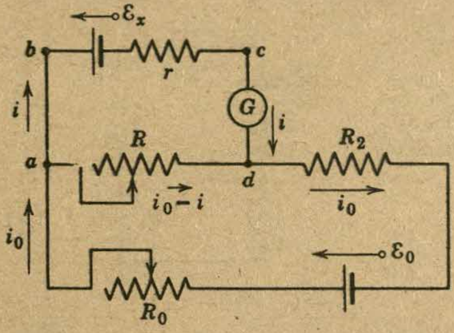


Fig. 32-9 Elements of a potentiometer

The emf can be obtained from Eq. 32–10 if the current i_0 is known. However, it is standard practice to replace $\mathcal{E}_x$ by a known standard emf $\mathcal{E}_s$, and once again to adjust R to the zero-current condition. This yields, assuming the current i_0 remains unchanged,

$$i_0 R_s = \mathcal{E}_s.$$

Combining the last two equations yields

$$\mathcal{E}_x = \mathcal{E}_s \frac{R_x}{R_s}, \tag{32-11}$$

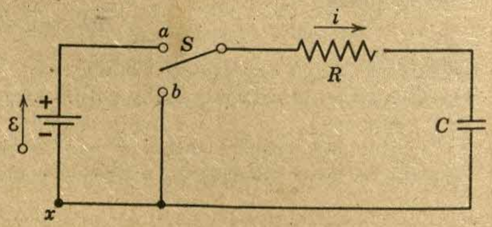
which allows us to compare emfs with precision. Note that the internal resistance r of the emf plays no role. In practice, potentiometers are conveniently packaged units, containing a *standard cell* which, after calibration at the National Bureau of Standards or elsewhere, serves as a convenient known standard seat of emf $\mathcal{E}_s$. Switching arrangements for replacing the unknown emf by the standard and arrangements for ascertaining that the current i_0 remains constant are also incorporated.

32–8 *RC* Circuits

The preceding sections dealt with circuits in which the circuit elements were resistors and in which the currents did not vary with time. Here we introduce the capacitor as a circuit element, which will lead us to the concept of time-varying currents. In Fig. 32–10 let switch S be thrown to position a. What current is set up in the single-loop circuit so formed? Let us apply conservation of energy principles.

In time dt a charge dq $(= i\,dt)$ moves through any cross section of the circuit. The work done by the seat of emf $(= \mathcal{E}\,dq$; see Eq. 32–1) must equal the energy that appears as Joule heat in the resistor during time

Fig. 32-10 An *RC* circuit.



dt ($= i^2R\,dt$) plus the increase in the amount of energy U that is stored in the capacitor [$= dU = d(q^2/2C)$; see Eq. 30-25]. In equation form

$$\mathcal{E}\,dq = i^2R\,dt + d\left(\frac{q^2}{2C}\right)$$

or

$$\mathcal{E}\,dq = i^2R\,dt + \frac{q}{C}\,dq.$$

Dividing by dt yields

$$\mathcal{E}\frac{dq}{dt} = i^2R + \frac{q}{C}\frac{dq}{dt}.$$

But dq/dt is simply i, so that this equation becomes

$$\mathcal{E} = iR + \frac{q}{C}. \qquad (32\text{-}12)$$

This equation also follows from the loop theorem, as it must, since the loop theorem was derived from the conservation of energy principle. Starting from point x and traversing the circuit clockwise, we experience an increase in potential in going through the seat of emf and decreases in potential in traversing the resistor and the capacitor, or

$$\mathcal{E} - iR - \frac{q}{C} = 0,$$

which is identical with Eq. 32-12.

We cannot immediately solve Eq. 32-12 because it contains two variables, q and i, which, however, are related by

$$i = \frac{dq}{dt}. \qquad (32\text{-}13)$$

Substituting for i into Eq. 32-12 gives

$$\mathcal{E} = R\frac{dq}{dt} + \frac{q}{C}. \qquad (32\text{-}14)$$

Our task now is to find the function $q(t)$ that satisfies this *differential equation.* Although this particular equation is not difficult to solve, we choose to avoid mathematical complexity by simply presenting the solution, which is

$$q = C\mathcal{E}(1 - e^{-t/RC}). \qquad (32\text{-}15)$$

We can easily test whether this function $q(t)$ is really a solution of Eq. 32-14 by substituting it into that equation and seeing whether an identity results. Differentiating Eq. 32-15 with respect to time yields

$$\frac{dq}{dt}(= i) = \frac{\mathcal{E}}{R}e^{-t/RC}. \qquad (32\text{-}16)$$

Substituting q (Eq. 32-15) and dq/dt (Eq. 32-16) into Eq. 32-14 yields an identity, as the student should verify. Thus Eq. 32-15 is a solution of Eq. 32-14.

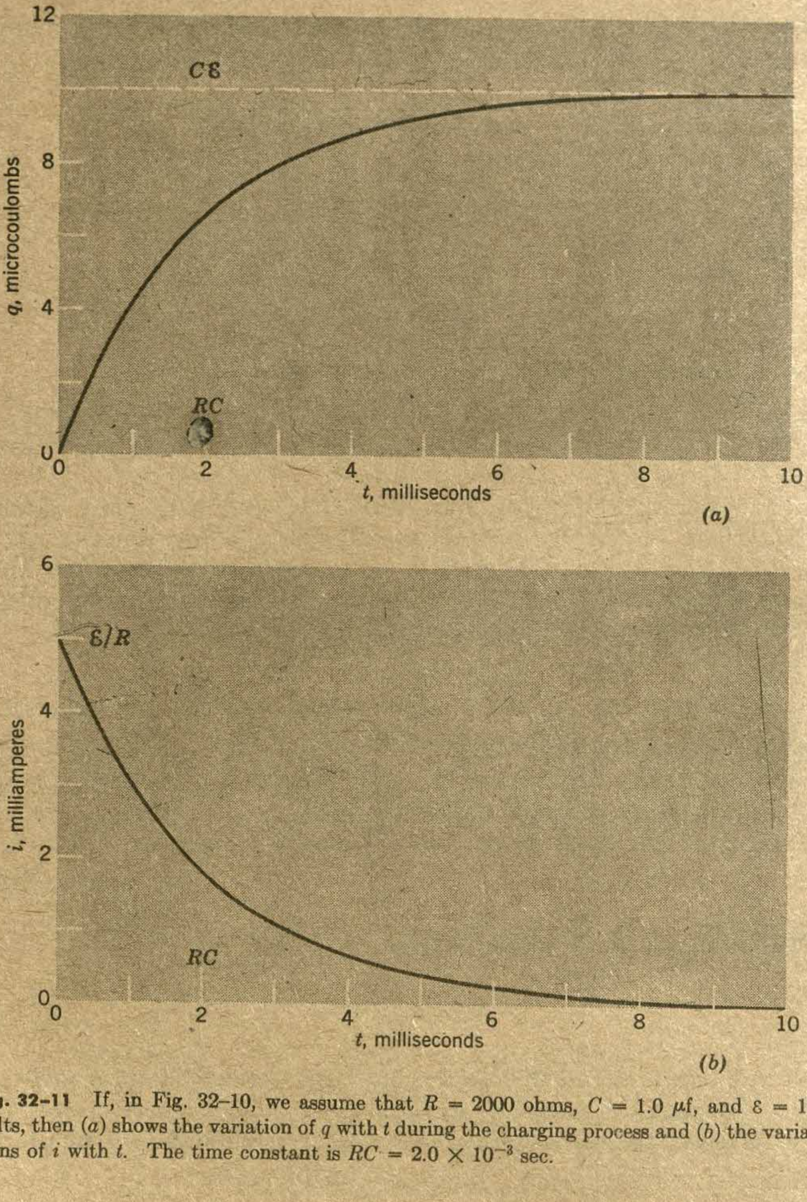


Fig. 32-11 If, in Fig. 32–10, we assume that $R = 2000$ ohms, $C = 1.0\ \mu$f, and $\mathcal{E} = 10$ volts, then (*a*) shows the variation of q with t during the charging process and (*b*) the variations of i with t. The time constant is $RC = 2.0 \times 10^{-3}$ sec.

Figure 32–11 shows some plots of Eqs. 32–15 and 32–16 for a particular case. Study of these plots and of the corresponding equations shows that (*a*) at $t = 0$, $q = 0$ and $i = \mathcal{E}/R$, and (*b*) as $t \to \infty$, $q \to C\mathcal{E}$ and $i \to 0$: that is, the current is initially $\mathcal{E}/R$ and finally zero; the charge on the capacitor plates is initially zero and finally $C\mathcal{E}$.

The quantity RC in Eqs. 32–15 and 32–16 has the dimensions of time (since the exponent must be dimensionless) and is called the *capacitative time constant* of the circuit. It is the time at which the charge on the capacitor has increased to within a factor of $(1 - e^{-1})$ $(= 63\%)$ of its equilibrium value. To show this, we put $t = RC$ in Eq. 32–15 to obtain

$$q = C\mathcal{E}(1 - e^{-1}) = 0.63C\mathcal{E}.$$

Since $C\mathcal{E}$ is the equilibrium charge on the capacitor, corresponding to $t \to \infty$, the foregoing statement follows.

▶ **Example 5.** After how many time constants will the energy stored in the capacitor in Fig. 32–10 reach one-half its equilibrium value?

The energy is given by Eq. 30–25, or

$$U = \frac{1}{2C}q^2,$$

the equilibrium energy U_∞ being $(1/2C)(C\mathcal{E})^2$. From Eq. 32–15, we can write the energy as

$$U = \frac{1}{2C}(C\mathcal{E})^2(1 - e^{-t/RC})^2$$

or

$$U = U_\infty(1 - e^{-t/RC})^2.$$

Putting $U = \frac{1}{2}U_\infty$ yields

$$\tfrac{1}{2} = (1 - e^{-t/RC})^2$$

and solving this relation for t yields finally

$$t = 1.22\,RC = 1.22 \text{ time constants.}$$ ◀

Figure 32–11 shows that if a resistance is included in the circuit the rate of increase of the charge of a capacitor toward its final equilibrium value is *delayed* in a way measured by the time constant RC. With no resistor present ($RC = 0$), the charge would rise immediately to its equilibrium value. Although we have shown that this time delay follows from an application of the loop theorem to RC circuits, it is important that the student develop a physical understanding of the causes of the delay.

When switch S in Fig. 32–10 is closed on a, the resistor experiences instantaneously an applied potential difference of $\mathcal{E}$, and an initial current of $\mathcal{E}/R$ is set up. Initially, the capacitor experiences no potential difference because its initial charge is zero, the potential difference always being given by q/C. The flow of charge through the resistor starts to charge the capacitor, which has several effects. First, the existence of a capacitor charge means that there must now be a potential difference $(= q/C)$ across the capacitor; this, in turn, means that the potential difference across the resistor must decrease by this amount, since the sum of the two potential differences must always equal $\mathcal{E}$. This decrease in the potential difference across R means that the charging current is reduced. Thus the charge of the capacitor builds up and the charging current decreases until the capacitor is fully charged. At this point the full emf $\mathcal{E}$ is applied to the capacitor, there being no potential drop ($i = 0$) across the resistor. This is precisely the reverse of the initial situation. The student should review the derivations of Eqs. 32–15 and 32–16 and should study Fig. 32–11 with the qualitative arguments of this paragraph in mind.

Assume now that the switch S in Fig. 32–10 has been in position a for a time t such that $t \gg RC$. The capacitor is then fully charged for all practical purposes. The switch S is then thrown to position b. How do the charge of the capacitor and the current vary with time?

With the switch S closed on b, there is no emf in the circuit and Eq. 32–12 for the circuit, with $\mathcal{E} = 0$, becomes simply

$$iR + \frac{q}{C} = 0. \tag{32–17}$$

Putting $i = dq/dt$ allows us to write, as the differential equation of the circuit (compare Eq. 32–14),

$$R\frac{dq}{dt} + \frac{q}{C} = 0. \tag{32–18a}$$

The solution is

$$q = q_0e^{-t/RC}, \tag{32–18b}$$

as the student may readily verify by substitution, q_0 being the initial charge on the capacitor. The capacitative time constant RC appears in this expression for capacitor discharge as well as in that for the charging process (Eq. 32–15). We see that at a time such that $t = RC$ the capacitor charge is reduced to q_0e^{-1}, which is 37% of the initial charge q_0.

The current during discharge follows from differentiating Eq. 32–18b, or

$$i = \frac{dq}{dt} = -\frac{q_0}{RC}e^{-t/RC} \tag{32–19}$$

The negative sign shows that the current is in the direction opposite to that shown in Fig. 32–10. This is as it should be, since the capacitor is discharging rather than charging. Since $q_0 = C\mathcal{E}$, we can write Eq. 32–19 as

$$i = -\frac{\mathcal{E}}{R}e^{-t/RC},$$

in which $\mathcal{E}/R$ appears as the initial current, corresponding to $t = 0$. This is reasonable because the initial potential difference for the fully charged capacitor is $\mathcal{E}$.

The behavior of the RC circuit of Fig. 32–10 during charge and discharge can be studied with a cathode-ray oscilloscope. This familiar laboratory device can display on its fluorescent screen plots of the variation of potential with time. Figure 32–12 shows the circuit of Fig. 32–10 with connections made to display (*a*) the potential difference V_C across the capacitor and (*b*) the potential difference V_R across the resistor as functions of time. V_C and V_R are given by

$$V_C = \left(\frac{1}{C}\right)q$$

and

$$V_R = (R)i,$$

the former being proportional to the charge and the latter to the current.

Figure 32–13 shows oscillograph plots of V_C and V_R that result when, in effect, switch S in Fig. 32–10 is thrown regularly back and forth between positions a and b, being left in each position for a time equal to several time constants. Intervals during which the charge is building up are labeled *ch* and those during which it is decaying are labeled *dis*.

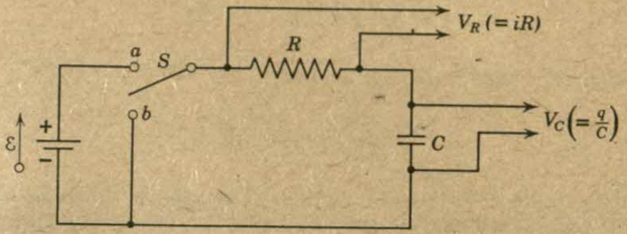


Fig. 32-12 The circuit of Fig. 32-10 with connections made to display the potential variations across the resistor and the capacitor on a cathode ray oscilloscope.

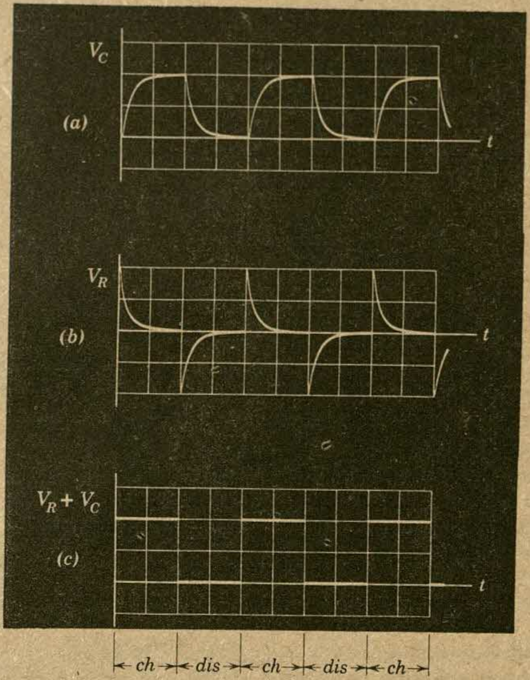


Fig. 32-13 In Fig. 32-10 switch S is thrown periodically, by electronic means, between positions a and b. The variations with time of the potential differences across (a) the capacitor, and (b) the resistor are shown, as displayed on a cathode ray oscilloscope. (c) The appearance of the screen when the oscilloscope is connected to display the sum of V_R and V_C. (Courtesy E. K. Hege, Rensselaer Polytechnic Institute.)

The "charge" intervals in plot *a* (see Eq. 32–15) are represented by

$$V = \left(\frac{1}{C}\right) q = \mathcal{E}(1 - e^{-t/RC})$$

and the "discharge" intervals (see Eq. 32–18*b*) by

$$V = \left(\frac{1}{C}\right) q = \mathcal{E}e^{-t/RC}.$$

Note that the current, as indicated by plot *b*, is in opposite directions during the charge and discharge intervals, in agreement with Eqs. 32–16 and 32–19.

In plot *c* in Fig. 32–13 the oscillograph has been connected to show the algebraic sum of plots *a* and *b*. According to the loop theorem this sum should equal $\mathcal{E}$ during the charge intervals and should be zero during the discharge intervals, when the battery is no longer in the circuit, that is,

$$V_R + V_C = \mathcal{E} \text{ during charge (see Eq. 32–12)}$$

$$V_R + V_C = 0 \text{ during discharge (see Eq. 32–17).}$$

Plot *c* is in exact agreement with this expectation.

QUESTIONS

1. Does the direction of the emf provided by a battery depend on the direction of current flow through the battery?

2. Figure 32–1*b* shows a gravitational analog of a single-loop electric circuit. Is the source of "gravitational emf" in this figure reversible as far as energy exchanges are concerned?

3. Discuss in detail the statement that the energy method and the loop theorem method for solving circuits are perfectly equivalent.

4. It is possible to generate a 10,000-volt potential difference by rubbing a pocket comb with wool. Why is this large voltage not dangerous when the much lower voltage provided by an ordinary electric outlet is very dangerous?

5. Devise a method for measuring the emf and the internal resistance of a battery.

6. A 25-watt, 110-volt bulb glows at normal brightness when connected across a bank of batteries. A 500-watt, 110-volt bulb glows only dimly when connected across the same bank. Explain.

7. Under what circumstances can the terminal potential difference of a battery exceed its emf?

8. What is the difference between an emf and a potential difference?

9. In the flow of incompressible fluids, what are the analogies to (*a*) the loop theorem and (*b*) the junction theorem?

10. Compare and contrast the formulas for the effective values of (*a*) capacitors and (*b*) resistors, in series and in parallel.

11. Does the time required for the charge on a capacitor in an *RC* circuit to build up to a given fraction of its equilibrium value depend on the value of the applied emf?

12. Devise a method whereby an *RC* circuit can be used to measure very high resistances.

PROBLEMS

1. A 5.0-amp current is set up in an external circuit by a 6.0-volt storage battery for 6.0 min. By how much is the chemical energy of the battery reduced?

2. The current in a simple series circuit is 5 amp. When an additional resistance of 2 ohms is inserted, the current drops to 4 amp. What was the resistance of the original circuit?

3. In Example 2 an ammeter whose resistance is 0.05 ohm is inserted in the circuit. What per cent change in the current results because of the presence of the meter?

4. In Fig. 32–3*a* put $\mathcal{E} = 2.0$ volts and $r = 100$ ohms. Plot (*a*) the current, and (*b*) the potential difference across R, as functions of R over the range 0 to 500 ohms. Make both plots on the same graph. (*c*) Make a third plot by multiplying together, for each value of R, the two curves plotted. What is the physical significance of this plot?

5. (*a*) In the circuit of Fig. 32–3*a* show that the power delivered to R as Joule heat is a maximum when R is equal to the internal resistance r of the battery. (*b*) Show that this maximum power is $P = \mathcal{E}^2/4r$.

6. Heat is to be generated in a 0.10-ohm resistor at the rate of 10 watts by connecting it to a battery whose emf is 1.5 volts. (*a*) What is the internal resistance of the battery? (*b*) What potential difference exists across the resistor?

7. (*a*) In Fig. 32–14 what value must R have if the current in the circuit is to be 0.001 amp? Take $\mathcal{E}_1 = 2.0$ volts, $\mathcal{E}_2 = 3.0$ volts, and $r_1 = r_2 = 3.0$ ohms. (*b*) What is the rate of Joule heating in R?

8. A wire of resistance 5.0 ohms is connected to a battery whose emf $\mathcal{E}$ is 2.0 volts and whose internal resistance is 1.0 ohm. In 2.0 min (*a*) how much energy is transferred from chemical to electric form? (*b*) How much energy appears in the wire as Joule heat? (*c*) Account for the difference between (*a*) and (*b*).

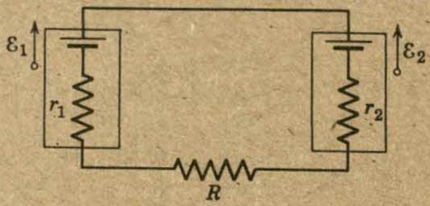


Fig. 32–14

9. In Fig. 32–6 calculate the potential difference between points c and d by as many paths as possible. Assume that $\mathcal{E}_1 = 4.0$ volts, $\mathcal{E}_2 = 1.0$ volt, $R_1 = R_2 = 10$ ohms, and $R_3 = 5$ ohms.

10. In Fig. 32–5 calculate the potential difference between a and c by considering a path that contains R and $\mathcal{E}_2$.

11. (*a*) In Fig. 32–15 what is the equivalent resistance of the network shown? (*b*) What are the currents in each resistor? Put $R_1 = 100$ ohms, $R_2 = R_3 = 50$ ohms, $R_4 = 75$ ohms, and $\mathcal{E} = 6.0$ volts.

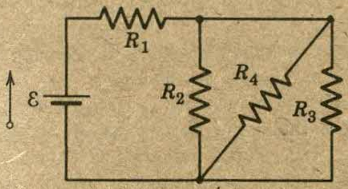


Fig. 32–15

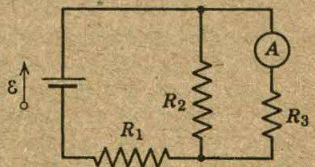


Fig. 32–16

12. In Fig. 32–16 imagine an ammeter inserted in the branch containing R_3. (*a*) What will it read, assuming $\mathcal{E} = 5.0$ volts, $R_1 = 2.0$ ohms, $R_2 = 4.0$ ohms, and $R_3 = 6.0$ ohms? (*b*) The ammeter and the source of emf are now physically interchanged. Show that the

ammeter reading remains unchanged. This reciprocity relationship holds for any circuit that contains only one source of emf.

13. Two batteries of emf $\mathcal{E}$ and internal resistance r are connected in parallel across a resistor R as in Fig. 32–20*b*. (*a*) For what value of R is the power delivered to the resistor a maximum? (*b*) What is the maximum power?

14. By using only two resistance coils—singly, in series, or in parallel—a student is able to obtain resistances of 3, 4, 12, and 16 ohms. What are the separate resistances of the coils?

15. *The Wheatstone bridge.* In Fig. 32–17 R_s is to be adjusted in value until points a and b are brought to exactly the same potential. (One tests for this condition by momentarily connecting a sensitive meter between a and b; if these points are at the same potential, the meter will not deflect.) Show that when this adjustment is made the following relation holds:

$$R_x = R_s \frac{R_2}{R_1}.$$

Unknown resistors (R_x) can be measured in terms of standards (R_s) using this device, which is called a Wheatstone bridge.

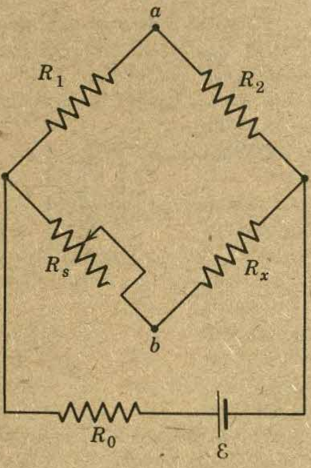


Fig. 32–17

16. If points a and b in Fig. 32–17 are connected by a wire of resistance r, show that the current in the wire is

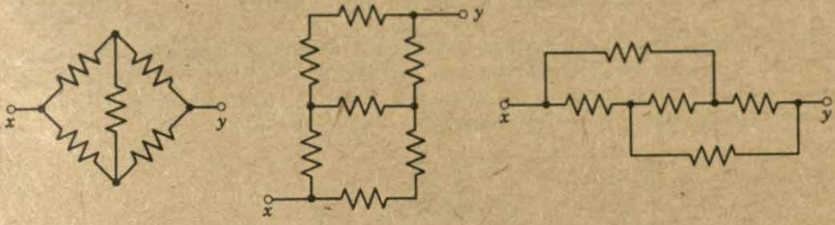
$$i = \frac{\mathcal{E}(R_s - R_x)}{(R + 2r)(R_s + R_x) + 2R_sR_x},$$

where $\mathcal{E}$ is the emf of the battery. Assume that R_1 and R_2 are equal ($R_1 = R_2 = R$) and that R_0 equals zero. Is this formula consistent with the result of Problem 15?

17. Two resistors, R_1 and R_2, may be connected either in series or parallel across a (resistanceless) battery with emf $\mathcal{E}$. We desire the Joule heating for the parallel combination to be five times that for the series combination. If R_1 equals 100 ohms, what is R_2?

18. Four 100-watt heating coils are to be connected in all possible series-parallel combinations and plugged into a 100-volt line. What different rates of heat dissipation are possible?

19. What is the equivalent resistance between the terminal points x and y of the circuits shown in (*a*) Fig. 32–18*a*, (*b*) Fig. 32–18*b*, and (*c*) Fig. 32–18*c*? Assume that the resistance of each resistor is 10 ohms.



(a) (b) (c)

Fig. 32-18

20. In Fig. 32-19 find the current in each resistor and the potential difference between a and b. Put $\mathcal{E}_1 = 6.0$ volts, $\mathcal{E}_2 = 5.0$ volts, $\mathcal{E}_3 = 4.0$ volts, $R_1 = 100$ ohms, and $R_2 = 50$ ohms.

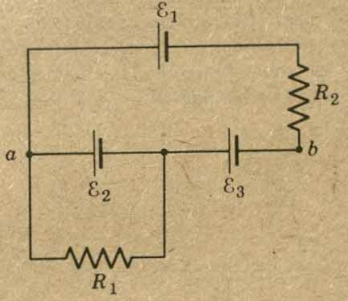
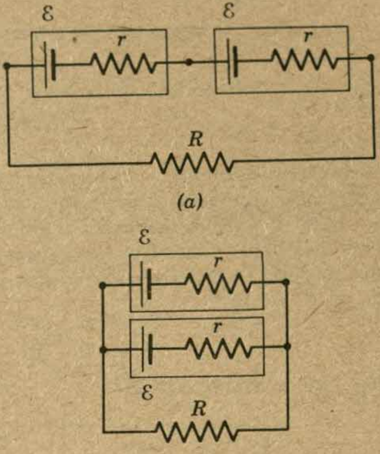


Fig. 32-19



(a)

(b)

Fig. 32-20

21. You are given two batteries of emf $\mathcal{E}$ and internal resistance r. They may be connected either in series or in parallel and are used to establish a current in a resistor R, as in Fig. 32-20. Derive expressions for the current in R for both methods of connection. Which connection yields the larger current if (a) $R > r$ and if (b) $R < r$?

22. (a) In Fig. 32-21 what power appears as Joule heat in R_1? In R_2? In R_3? (b) What power is supplied by $\mathcal{E}_1$? By $\mathcal{E}_2$? (c) Discuss the energy balance in this circuit. Assume that $\mathcal{E}_1 = 3.0$ volts, $\mathcal{E}_2 = 1.0$ volt, $R_1 = 5.0$ ohms, $R_2 = 2.0$ ohms, and $R_3 = 4.0$ ohms.

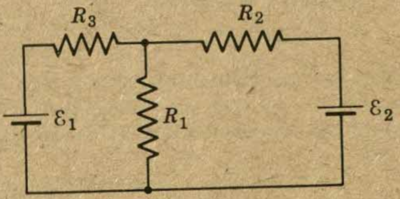


Fig. 32-21

23. For manual control of the current in a circuit, a student uses a parallel combination of variable resistors of the sliding contact type, as in Fig. 32–22, with $R_1 = 20R_2$. (*a*) What procedure is used to adjust the current to the desired value? (*b*) Why is the parallel combination better than a single-variable resistor?

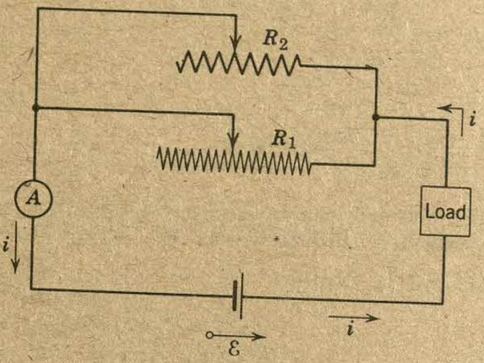


Fig. 32–22

24. *Resistance measurement.* A voltmeter (resistance R_V) and an ammeter (resistance R_A) are connected to measure a resistance R, as in Fig. 32–23*a*. The resistance is given by $R = V/i$, where V is the voltmeter reading and i is the current *in the resistor R*. Some of the current registered by the ammeter (i') goes through the voltmeter so that the ratio of the meter readings ($= V/i'$) gives only an *apparent* resistance reading R'. Show that R and R' are related by

$$\frac{1}{R} = \frac{1}{R'} - \frac{1}{R_V}.$$

Note that if $R_V \gg R$, then $R \cong R'$.

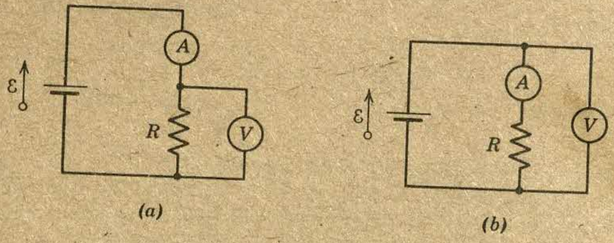


Fig. 32–23

25. *Resistance measurement.* If meters are used to measure resistance, they may also be connected as they are in Fig. 32–23*b*. Again the ratio of the meter readings gives only an apparent resistance R'. Show that R' is related to R by

$$R = R' - R_A,$$

in which R_A is the ammeter resistance. Note that if $R_A \ll R$, then $R \cong R'$.

26. In Fig. 32–8 assume that $\mathcal{E} = 5.0$ volts, $r = 2.0$ ohms, $R_1 = 5.0$ ohms, and $R_2 = 4.0$ ohms. If $R_A = 0.10$ ohm, what per cent error is made in reading the current? Assume that the voltmeter is not present.

27. In Fig. 32–8 assume that $\mathcal{E} = 5.0$ volts, $r = 20$ ohms, $R_1 = 50$ ohms, and $R_2 = 40$ ohms. If $R_v = 1000$ ohms, what per cent error is made in reading the potential differences across R_1? Ignore the presence of the ammeter.

28. (*a*) Find the three currents in Fig. 32–24. (*b*) Find V_{ab}. Assume that $R_1 = 1.0$ ohm, $R_2 = 2.0$ ohms, $\mathcal{E}_1 = 2.0$ volts, and $\mathcal{E}_2 = \mathcal{E}_3 = 4.0$ volts.

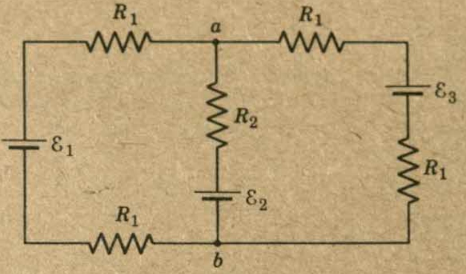


Fig. 32–24

29. How many time constants must elapse before a capacitor in an *RC* circuit is charged to within 1.0 per cent of its equilibrium charge?

30. In the circuit of Fig. 32–25 let i_1, i_2, and i_3 be the currents through resistors R_1, R_2, and R_3, respectively, and let V_1, V_2, V_3, and V_C be the corresponding potential differences across the resistors and across the capacitor *C*. (*a*) Plot qualitatively as a function of time after switch *S* is closed the currents and voltages listed above. (*b*) After being closed for a large number of time constants, the switch *S* is now opened. Plot qualitatively as a function of time after the switch is opened the currents and voltages listed above.

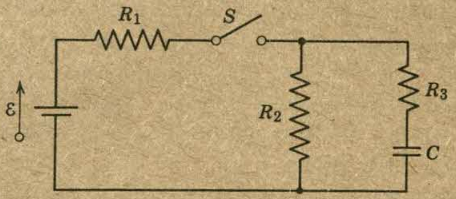


Fig. 32–25

31. Show that the units of *RC* are indeed time units, that is, that 1 ohm × 1 farad = 1 sec.

32. Prove that when switch *S* in Fig. 32–10 is thrown from *a* to *b* all the energy stored in the capacitor is transformed into Joule heat in the resistor. Assume that the capacitor is fully charged before the switch is thrown.

33. A 3.0×10^6-ohm resistor and a 1.0-μf capacitor are connected in a single-loop circuit with a seat of emf with $\mathcal{E} = 4.0$ volts. At 1.0 sec after the connection is made, what are the rates at which (*a*) the charge of the capacitor is increasing, (*b*) energy is being stored in the capacitor, (*c*) Joule heat is appearing in the resistor, and (*d*) energy is being delivered by the seat of emf?

The Magnetic Field

CHAPTER 33

33-1 The Magnetic Field

The science of magnetism grew from the observation that certain "stones" (magnetite) would attract bits of iron. The word *magnetism* comes from the district of Magnesia in Asia Minor, which is one of the places at which the stones were found. Figure 33-1 shows a modern permanent magnet, the lineal descendant of these natural magnets. Another "natural magnet" is the earth itself, whose orienting action on a magnetic compass needle has been known since ancient times.

In 1820 Oersted first discovered that a current in a wire can also produce magnetic effects, namely, that it can change the orientation of a compass needle. We pointed out in Section 26-1 how this important discovery linked the then separate sciences of magnetism and electricity. The magnetic ef-

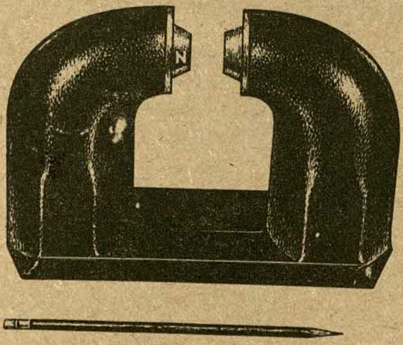


Fig. 33-1 A permanent magnet. Lines of magnetic induction leave the north pole face, marked N, and enter the south pole face on the other side of the air gap.

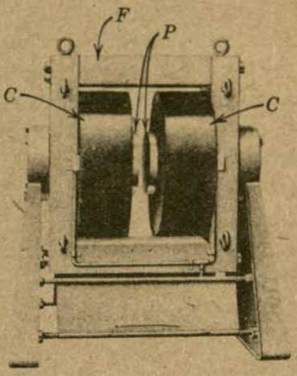


Fig. 33–2 A research-type electromagnet showing iron frame F, pole faces P, and coils C. The pole faces are 12 in. in diameter. (Courtesy Varian Associates.)

fect of a current in a wire can be intensified by forming the wire into a coil of many turns and by providing an iron core. Figure 33–2 shows how this is done in a large electromagnet of a type commonly used for research involving magnetism.

We define the space around a magnet or a current-carrying conductor as the site of a *magnetic field*, just as we defined the space near a charged rod as the site of an electric field. The basic magnetic field vector **B**, which we define in the following section, is called the *magnetic induction:*[*] it can be represented by *lines of induction*, just as the electric field was represented by lines of force. As for the electric field (see Section 27–3), the magnetic field vector is related to its lines of induction in this way:

1. The tangent to a line of induction at any point gives the *direction* of **B** at that point.
2. The lines of induction are drawn so that the number of lines per unit cross-sectional area is proportional to the *magnitude* of the magnetic field vector **B**. Where the lines are close together B is large and where they are far apart B is small.

As for the electric field, the field vector **B** is of fundamental importance, the lines of induction simply giving a graphic representation of the way **B** varies throughout a certain region of space.

The *flux* Φ_B for a magnetic field can be defined in exact analogy with the flux Φ_E for the electric field, namely

$$\Phi_B = \int \mathbf{B} \cdot d\mathbf{S}, \tag{33–1}$$

in which the integral is taken over the surface (closed or open) for which Φ_B is defined.

---

[*] *Magnetic field strength* would be a more suitable name for **B**, but it has been usurped for historical reasons by another vector connected with the magnetic field (see Section 37–7).

33-2 The Definition of B

This chapter is not concerned with the *causes* of the magnetic field; we seek to determine (*a*) whether a magnetic field exists at a given point and (*b*) the action of this field on charges moving through it. As for the electric field, a particle of charge q_0 serves as a test body. We assume that there is no electric field present, which means that, neglecting gravity, no force will act on the test body if it is placed *at rest* at the point in question.

Let us fire a positive test charge with arbitrary velocity **v** through a point P. *If a sideways deflecting force* **F** *acts on it, we assert that a magnetic field is present at* P and we define the magnetic induction **B** of this field in terms of **F** and other measured quantities.

If we vary the direction of **v** through point P, keeping the magnitude of **v** unchanged, we find, in general, that although **F** will always remain at right angles to **v** its magnitude F will change. For a particular orientation of **v** (and also for the opposite orientation $-\mathbf{v}$) the force **F** becomes zero. We define this direction as the direction of **B**, the specification of the sense of **B** (that is, the way it points along this line) being left to the more complete definition of **B** that we give below.

Having found the *direction* of **B**, we are now able to orient **v** so that the test charge moves at right angles to **B**. We will find that the force **F** is now a maximum, and we define the *magnitude* of **B** from the measured magnitude of this maximum force $F_\perp$, or

$$B = \frac{F_\perp}{q_0 v}. \tag{33-2}$$

Let us regard this definition of **B** (in which we have specified its magnitude and direction, but not its sense) as preliminary to the complete vector definition that we now give: *If a positive test charge q_0 is fired with velocity* **v** *through a point P and if a (sideways) force* **F** *acts on the moving charge, a magnetic induction* **B** *is present at point P, where* **B** *is the vector that satisfies the relation*

$$\mathbf{F} = q_0\mathbf{v} \times \mathbf{B}, \tag{33-3a}$$

v, q_0, and **F** being measured quantities. The magnitude of the magnetic deflecting force **F**, according to the rules for vector products, is given by *

$$F = q_0 vB \sin\theta, \tag{33-3b}$$

in which θ is the angle between **v** and **B**.

Figure 33-3 shows the relations among the vectors. We see that **F**, being at right angles to the plane formed by **v** and **B**, will always be at right angles to **v** (and also to **B**) and thus will always be a sideways force. Equation 33-3*a* is consistent with the observed facts that (*a*) the magnetic force vanishes as $v \to 0$, (*b*) the magnetic force vanishes if **v** is either parallel or antiparallel to the direction of **B** (in these cases $\theta = 0$ or $180°$ and $\mathbf{v} \times \mathbf{B} = 0$), and (*c*) if **v** is at right angles to **B** ($\theta = 90°$), the deflecting force has its maximum value, given by Eq. 33-2, that is, $F_\perp = q_0 vB$.

---

* The student may wish to review Section 2-4, which deals with vector products.

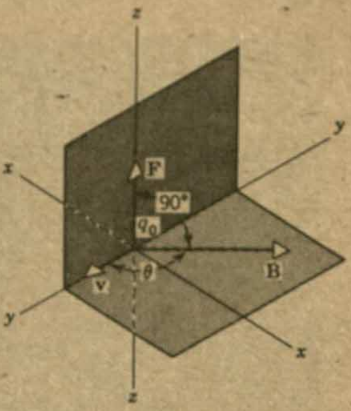


Fig. 33-3 Illustrating $\mathbf{F} = q_0\mathbf{v} \times \mathbf{B}$ (Eq. 33-3a).

This definition of **B** is similar in spirit, although more complex, than the definition of the electric field strength **E**, which we can cast into this form: *If a positive test charge q_0 is placed at point P and if an (electric) force* **F** *acts on the stationary charge, an electric field* **E** *is present at P, where* **E** *is the vector satisfying the relation*

$$\mathbf{F} = q_0\mathbf{E},$$

q_0 and **F** being measured quantities. In defining **E**, the only characteristic direction to appear is that of the electric force $\mathbf{F}_E$ which acts on the positive test body; the direction of **E** is taken to be that of $\mathbf{F}_E$. In defining **B**, *two* characteristic directions appear, those of **v** and of the magnetic force $\mathbf{F}_B$; they prove always to be at right angles.

In Fig. 33-4 a positive and a negative electron are created at point *P* in a bubble chamber. A magnetic field is perpendicular to the chamber, pointing out of the plane of the figure (symbol ⊙).* The relation $\mathbf{F} = q_0\mathbf{v} \times \mathbf{B}$ (Eq. 33-3a) shows that the deflecting forces acting on the two particles are as indicated in the figure. These deflecting forces would make the tracks deflect as shown.

The unit of **B** that follows from Eq. 33-3 is the (nt/coul)/(meter/sec). This is given the special name weber/meter², or tesla. Recalling that a coul/sec is an ampere,

$$1 \text{ weber/meter}^2 = \frac{1 \text{ nt}}{\text{coul (meter/sec)}} = \frac{1 \text{ nt}}{\text{amp-m}}$$

An earlier unit for **B**, still in common use, is the *gauss;* the relationship is

$$1 \text{ weber/meter}^2 = 10^4 \text{ gauss.}$$

The *weber* is used to measure Φ_B, the flux of **B**; see Eq. 33-1.

The fact that the magnetic force is always at right angles to the direction of motion means that (for steady magnetic fields) the work done by this

---

* The symbol ⊗ indicates a vector into the page, the × being thought of as the tail of an arrow; the symbol ⊙ indicates a vector out of the page, the dot being thought of as the tip of an arrow.

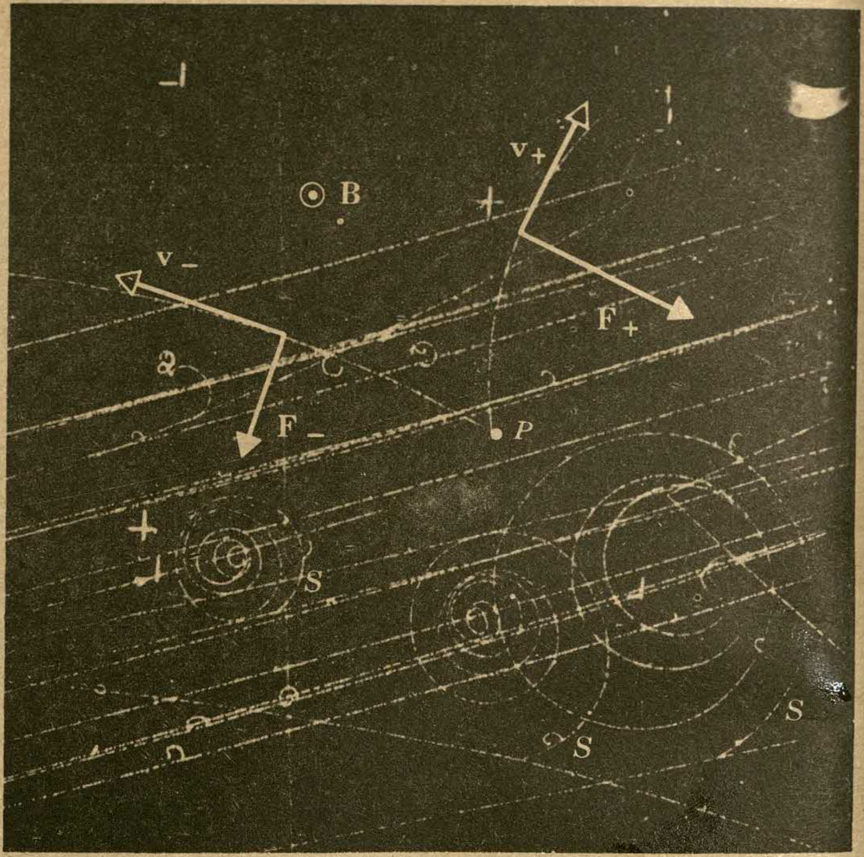


Fig. 33–4 A *bubble chamber* is a device for rendering visible, by means of small bubbles, the tracks of charged particles that pass through the chamber. The figure is a photograph taken with such a chamber immersed in a field of magnetic induction B and exposed to radiations from a large cyclotron-like accelerator. The curved V at point P is formed by a positive and a negative electron, which deflect in opposite directions in the magnetic field. The spirals S are the tracks of three low-energy electrons. (Courtesy E. O. Lawrence Radiation Laboratory, University of California.)

force on the particle is zero. For an element of the path of the particle of length $d\mathbf{l}$, this work dW is $\mathbf{F}_B \cdot d\mathbf{l}$; dW is zero because $\mathbf{F}_B$ and $d\mathbf{l}$ are always at right angles. Thus a static magnetic field cannot change the kinetic energy of a moving charge; it can only deflect it sideways.

If a charged particle moves through a region in which both an electric field and a magnetic field are present, the resultant force is found by combining Eqs. 27–2 and 33–3*a*, or

$$\mathbf{F} = q_0\mathbf{E} + q_0\mathbf{v} \times \mathbf{B}. \tag{33–4}$$

This is sometimes called the *Lorentz relation* in tribute to H. A. Lorentz who did so much to develop and clarify the concepts of the electric and magnetic fields.

▶ **Example 1.** A uniform field of magnetic induction **B** points horizontally from south to north; its magnitude is 1.5 webers/meter². If a 5.0-Mev proton moves vertically downward through this field, what force will act on it?

The kinetic energy of the proton is

$$K = (5.0 \times 10^6 \text{ ev})(1.6 \times 10^{-19} \text{ joule/ev}) = 8.0 \times 10^{-13} \text{ joule}.$$

Its speed can be found from the relation $K = \frac{1}{2}mv^2$, or

$$v = \sqrt{\frac{2K}{m}} = \sqrt{\frac{(2)(8.0 \times 10^{-13} \text{ joule})}{1.7 \times 10^{-27} \text{ kg}}} = 3.1 \times 10^7 \text{ meters/sec}.$$

Equation 33-3*b* gives

$$F = qvB \sin\theta = (1.6 \times 10^{-19} \text{ coul})(3.1 \times 10^7 \text{ meters/sec})(1.5 \text{ webers/meter}^2)(\sin 90°)$$
$$= 7.4 \times 10^{-12} \text{ nt}.$$

The student can show that this force is about 4×10^{14} times greater than the weight of the proton.

The relation $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$ shows that the *direction* of the deflecting force is to the east. If the particle had been negatively charged, the deflection would have been to the west. This is predicted automatically by Eq. 33-3*a* if we substitute $-e$ for q. ◀

33-3 Magnetic Force on a Current

A current is an assembly of moving charges. Because a magnetic field exerts a sideways force on a moving charge, we expect that it will also exert a sideways force on a wire carrying a current. Figure 33-5 shows a length l of wire carrying a current i and placed in a field of magnetic induction **B**. For simplicity we have oriented the wire so that the current density vector **j** is at right angles to **B**.

The current i in a metal wire is carried by the free (or conduction) electrons, n being the number of such electrons per unit volume of the wire. The magnitude of the average force on one such electron is given by Eq. 33-3*b*, or, since $\theta = 90°$,

$$F' = q_0 vB \sin\theta = ev_d B$$

where v_d is the drift speed. From the relation $v_d = j/ne$ (Eq. 31-5),

$$F' = e\left(\frac{j}{ne}\right)B = \frac{jB}{n}.$$

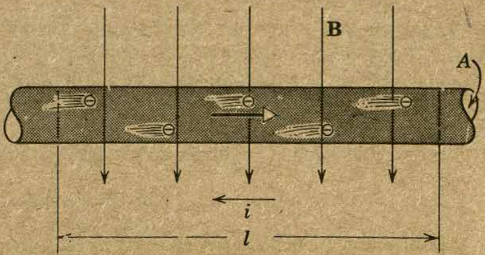


Fig. 33-5 A wire carrying a current i is placed at right angles to a field of magnetic induction **B**.

The length l of the wire contains nAl free electrons, Al being the volume of the section of wire of cross section A that we are considering. The total force on the free electrons in the wire, and thus on the wire itself, is

$$F = (nAl)F' = nAl\frac{jB}{n}.$$

Since jA is the current i in the wire, we have

$$F = ilB. \tag{33-5}$$

The negative charges, which move to the right in the wire of Fig. 33-5, are equivalent to positive charges moving to the left, that is, in the direction of the current arrow. For such a positive charge the velocity $\mathbf{v}$ would point to the left and the force on the wire, given by Eq. 33-3a ($\mathbf{F} = q_0\mathbf{v} \times \mathbf{B}$) points up, out of the page. This same conclusion follows if we consider the actual negative charge carriers for which $\mathbf{v}$ points to the right but q_0 has a negative sign. Thus by measuring the sideways magnetic force on a wire carrying a current and placed in a magnetic field we cannot tell whether the current carriers are negative charges moving in a given direction or positive charges moving in the opposite direction.

Equation 33-5 holds only if the wire is at right angles to $\mathbf{B}$. We can express the more general situation in vector form as

$$\mathbf{F} = i\mathbf{l} \times \mathbf{B}, \tag{33-6a}$$

where $\mathbf{l}$ is a (displacement) vector that points along the (straight) wire in the direction of the current. Equation 33-6a is equivalent to the relation $\mathbf{F} = q_0\mathbf{v} \times \mathbf{B}$ (Eq. 33-3a); either can be taken as a defining equation for $\mathbf{B}$. The student should note that the vector $\mathbf{l}$ in Fig. 33-5 points to the left and that the magnetic force $\mathbf{F}$ ($= i\mathbf{l} \times \mathbf{B}$) points up, out of the page. This agrees with the conclusion obtained by analyzing the forces that act on the individual charge carriers.

If we consider a differential element of a conductor of length $d\mathbf{l}$, the force $d\mathbf{F}$ acting on it can be found, by analogy with Eq. 33-6a, from

$$d\mathbf{F} = i\, d\mathbf{l} \times \mathbf{B}. \tag{33-6b}$$

By integrating this formula in an appropriate way we can find the force $\mathbf{F}$ on a nonlinear conductor.

▶ **Example 2.** A wire bent as shown in Fig. 33-6 carries a current i and is placed in a uniform field of magnetic induction $\mathbf{B}$ that emerges from the plane of the figure. Calculate the force acting on the wire. The magnetic field is represented by lines of induction, shown emerging from the page. The dots show that the sense of $\mathbf{B}$ is up, out of the page.

The force on each straight section, from Eq. 33-6a, has the magnitude

$$F_1 = F_3 = ilB$$

and points down as shown by the arrows in the figure.

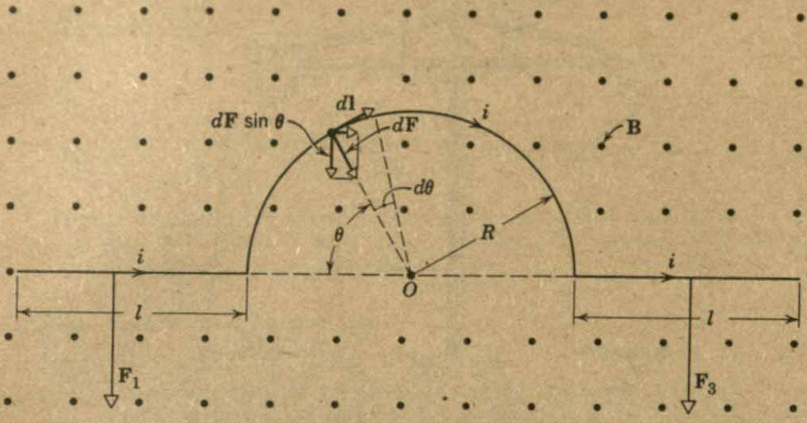


Fig. 33-6 Example 2.

A segment of wire of length dl on the arc has a force $d\mathbf{F}$ on it whose magnitude is

$$dF = iB\,dl = iB(R\,d\theta)$$

and whose direction is radially toward O, the center of the arc. Only the downward component of this force is effective, the horizontal component being canceled by an oppositely directed component associated with the corresponding arc segment on the other side of O. Thus the total force on the semicircle of wire about O points down and is

$$F_2 = \int_0^\pi dF \sin\theta = \int_0^\pi (iBR\,d\theta)\sin\theta = iBR\int_0^\pi \sin\theta\,d\theta = 2iBR.$$

The resultant force on the whole wire is

$$F = F_1 + F_2 + F_3 = 2ilB + 2iBR = 2iB(l + R).$$

Notice that this force is the same as that acting on a straight wire of length $2l + 2R$. ◀

Figure 33-7 shows the arrangement used by Thomas, Driscoll, and Hipple at the National Bureau of Standards to measure the magnetic induction provided by a laboratory magnet such as that of Fig. 33-2. The rectangle is a coil of nine turns whose width a and length b are about 10 cm and 70 cm, respectively. The lower end of the coil is placed in the field of magnetic induction **B** and the upper end is hung from the arm of a balance; **B** enters the plane of the figure at right angles.

An accurately known current i of about 0.10 amp is set up in the coil in the direction shown, and weights are placed in the right-hand balance pan until the system is balanced. The magnetic force $\mathbf{F}\ (= i\mathbf{l} \times \mathbf{B}$; see Eq. 33-6$a$) on the bottom leg of the coil points upward, as shown in the figure. Equation 33-5 also shows that the force *on each wire* at the bottom of the coil is iaB. Since there are nine wires, the total force on the bottom leg of the coil is $9iaB$. The forces on the vertical sides of the coil $(= i\mathbf{l} \times \mathbf{B})$ are sideways; because they are equal and opposite, they cancel and produce no effect.

After balancing the system, the experimenters reversed the direction of the current, which changed the sign of all the magnetic forces acting on the coil. In particular, **F** then pointed downward, which caused the rest point of the balance to move. A mass

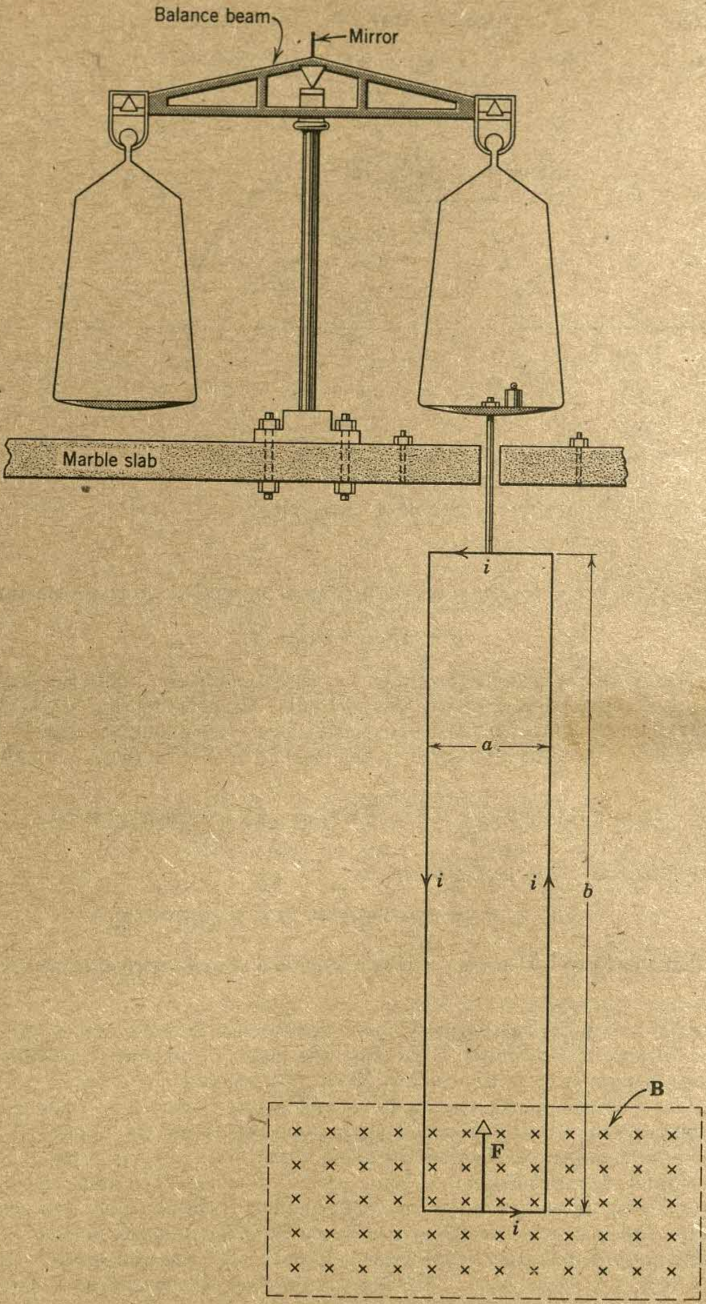


Fig. 33–7 Apparatus used to measure **B**. The zero point of the balance is observed by means of a light beam reflected from the mirror attached to the balance beam.

m of about 8.78 gm had to be added to the left balance pan to restore the original rest point. The *change* in force when the current is reversed is $2F$, and this must equal the weight added to the left balance pan, or

$$mg \doteq 2(9iaB) = 18iaB.$$

This gives

$$B = \frac{mg}{18ai} = \frac{(8.78 \times 10^{-3}\ \text{kg})(9.80\ \text{meters/sec}^2)}{(18)(0.10\ \text{meter})(0.10\ \text{amp})} = 0.48\ \text{weber/meter}^2 = 4800\ \text{gauss}.$$

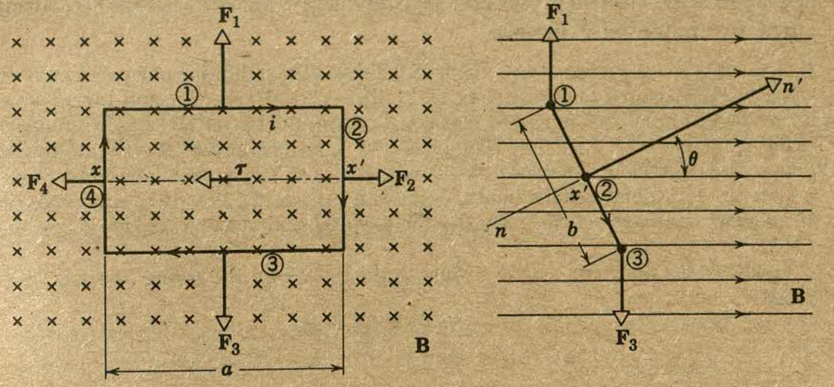
The Bureau of Standards' workers made this measurement with much more care than these approximate numbers suggest. In one series of measurements, for example, they found a magnetic induction of 4697.55 gauss.

33-4 Torque on a Current Loop

Figure 33-8 shows a rectangular loop of wire of length a and width b placed in a uniform field of induction **B**, with sides 1 and 3 always normal to the field direction. The normal nn' to the plane of the loop makes an angle θ with the direction of **B**.

Assume the current to be as shown in the figure. Wires must be provided to lead the current into the loop and out of it. If these wires are twisted tightly together, there will be no net magnetic force on the twisted pair because the currents in the two wires are in opposite directions. Thus the lead wires may be ignored. Also, some way of supporting the loop must be provided. Let us imagine it to be suspended from a long string attached to the loop at its center of mass. In this way the loop will be free to turn, through a small angle at least, about any axis through the center of mass.

The net force on the loop is the resultant of the forces on the four sides of the loop. On side 2 the vector **l** points in the direction of the current and has the magnitude b. The angle between **l** and **B** for side 2 (see Fig. 33-8b) is



(a) *(b)*

Fig. 33-8 A rectangular coil carrying a current i is placed in a uniform external magnetic field.

$90° - \theta$. Thus the magnitude of the force on this side is

$$F_2 = ibB \sin (90° - \theta) = ibB \cos \theta.$$

From the relation $\mathbf{F} = i\mathbf{l} \times \mathbf{B}$ (Eq. 33–6*a*), we find the direction of $\mathbf{F}_2$ to be out of the plane of Fig. 33–8*b*. The student can show that the force $\mathbf{F}_4$ on side 4 has the same magnitude as $\mathbf{F}_2$ but points in the opposite direction. Thus $\mathbf{F}_2$ and $\mathbf{F}_4$, taken together, have no effect on the motion of the loop. The net force they provide is zero, and, since they have the same line of action, the net torque due to these forces is also zero.

The common magnitude of $\mathbf{F}_1$ and $\mathbf{F}_3$ is iaB. These forces, too, are oppositely directed so that they do not tend to move the coil bodily. As Fig. 33–8*b* shows, however, they do *not* have the same line of action if the coil is in the position shown; there is a net torque, which tends to rotate the coil clockwise about the line xx. The coil can be supported on a rigid axis that lies along xx', with no loss of its freedom of motion. This torque can be represented in Fig. 33–8*b* by a vector pointing into the figure at point x' or in Fig. 33–8*a* by a vector pointing along the xx' axis from right to left.

The magnitude of the torque τ' is found by calculating the torque caused by $\mathbf{F}_1$ about axis xx' and doubling it, for $\mathbf{F}_3$ exerts the same torque about this axis that $\mathbf{F}_1$ does. Thus

$$\tau' = 2(iaB)\left(\frac{b}{2}\right)(\sin \theta) = iabB \sin \theta.$$

This torque acts on every turn of the coil. If there are N turns, the torque on the entire coil is

$$\tau = N\tau' = NiabB \sin \theta = NiAB \sin \theta, \tag{33–7}$$

in which A, the area of the coil, is substituted for ab.

This equation can be shown to hold for *all plane loops of area A, whether they are rectangular or not.* A torque on a current loop is the basic operating principle of the electric motor and of most electric meters used for measuring current or potential difference.

▶ **Example 3.** *A galvanometer.* Figure 33–9 shows the rudiments of a galvanometer, which is a device used to measure currents. The coil is 2.0 cm high and 1.0 cm wide; it has 250 turns and is mounted so that it can rotate about a vertical axis in a uniform *radial* magnetic field with $B = 2000$ gauss. A spring Sp provides a countertorque that cancels out the magnetic torque, resulting in a steady angular deflection ϕ corresponding to a given current i in the coil. If a current of 1.0×10^{-4} amp produces an angular deflection of 30°, what is the *torsional constant* κ of the spring (see Eq. 15–21)?

Equating the magnetic torque to the torque caused by the spring (see Eq. 33–7) yields

$$\tau = NiAB \sin \theta = \kappa\phi$$

or

$$\begin{aligned}\kappa &= \frac{NiAB \sin \theta}{\phi} \\ &= \frac{(250)(1.0 \times 10^{-4}\text{ amp})(2.0 \times 10^{-4}\text{ meter}^2)(0.20\text{ weber/meter}^2)(\sin 90°)}{30°} \\ &= 3.3 \times 10^{-8}\text{ nt-m/deg.}\end{aligned}$$

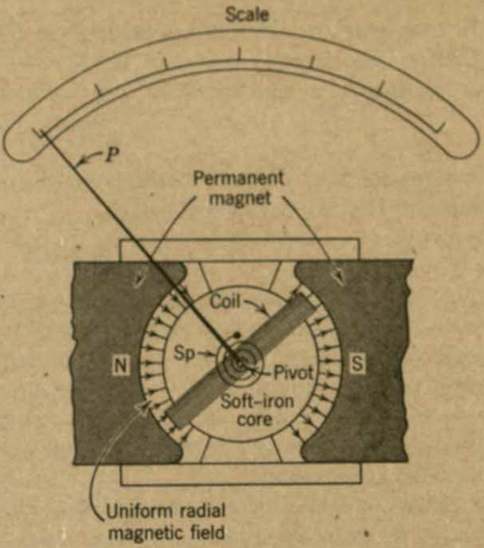


Fig. 33-9 Example 3. The elements of a galvanometer, showing the coil, the helical spring *Sp*, and pointer *P*.

Note that the normal to the plane of the coil (that is, the pointer *P*) is always at right angles to the (radial) magnetic field so that $\theta = 90°$. ◀

A current loop orienting itself in an external magnetic field reminds us of the action of a compass needle in such a field. One face of the loop behaves like the north pole of the needle; * the other face behaves like the south pole. Compass needles, bar magnets, and current loops can all be regarded as *magnetic dipoles*. We show this here for the current loop, reasoning entirely by analogy with *electric* dipoles.

A structure is called an electric dipole if (*a*) when placed in an *external* electric field it experiences a torque given by Eq. 27-11,

$$\boldsymbol{\tau} = \mathbf{p} \times \mathbf{E}, \tag{33-8}$$

where $\mathbf{p}$ is the electric dipole moment, and (*b*) it sets up a field of its own at distant points, described qualitatively by the lines of force of Fig. 29-10 and quantitatively by Eq. 29-11. These two requirements are not independent; if one is fulfilled, the other follows automatically.

The magnitude of the torque described by Eq. 33-8 is

$$\tau = pE \sin \theta, \tag{33-9}$$

where θ is the angle between $\mathbf{p}$ and $\mathbf{E}$. Let us compare this with Eq. 33-7, the expression for the torque on a current loop:

$$\tau = (NiA)B \sin \theta. \tag{33-7}$$

---

* The north pole of a compass needle is the end that points toward the geographic north.

In each case the appropriate field (E or B) appears, as does a term $\sin\theta$. Comparison suggests that NiA in Eq. 33–7 can be taken as the *magnetic dipole moment* μ, corresponding to p in Eq. 33–9, or

$$\mu = NiA. \qquad (33\text{–}10)$$

Equation 33–7 suggests that we write the torque on a current loop in vector form, in analogy with Eq. 33–8, or

$$\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B}. \qquad (33\text{–}11)$$

The magnetic dipole moment of the loop $\boldsymbol{\mu}$ must be taken to lie along the axis of the loop; its direction is given by the following rule: Let the fingers of the right hand curl around the loop in the direction of the current; the extended right thumb will then point in the direction of $\boldsymbol{\mu}$. The student should check carefully that, if $\boldsymbol{\mu}$ is defined by this rule and Eq. 33–10, Eq. 33–11 correctly describes in every detail the torque acting on a current loop in an external field (see Fig. 33–8).

Since a torque acts on a current loop, or other magnetic dipole, when it is placed in an external magnetic field, it follows that work (positive or negative) must be done by an external agent to change the orientation of such a dipole. Thus a magnetic dipole has *potential energy* associated with its orientation in an external magnetic field. This energy may be taken to be zero for any arbitrary position of the dipole. By analogy with the assumption made for electric dipoles in Section 27–6, we assume that the magnetic energy U is zero when $\boldsymbol{\mu}$ and $\mathbf{B}$ are at right angles, that is, when $\theta = 90°$. This choice of a zero-energy configuration for U is arbitrary because we are interested only in the *changes* in energy that occur when the dipole is rotated.

The magnetic potential energy in any position θ is defined as the work that an external agent must do to turn the dipole from its zero-energy position ($\theta = 90°$) to the given position θ. Thus

$$U = \int_{90°}^{\theta} \tau\, d\theta = \int_{90°}^{\theta} NiAB \sin\theta\, d\theta = \mu B \int_{90°}^{\theta} \sin\theta\, d\theta = -\mu B \cos\theta,$$

in which Eq. 33–7 is used to substitute for τ. In vector symbolism this relation can be written as

$$U = -\boldsymbol{\mu}\cdot\mathbf{B}, \qquad (33\text{–}12)$$

which is in perfect correspondence with Eq. 27–13, the expression for the energy of an *electric* dipole in an external *electric* field,

$$U = -\mathbf{p}\cdot\mathbf{E}.$$

▶ **Example 4.** A circular coil of N turns has an effective radius a and carries a current i. How much work is required to turn it in an external magnetic field $\mathbf{B}$ from a position in which θ equals zero to one in which θ equals 180°? Assume that $N = 100$, $a = 5.0$ cm, $i = 0.10$ amp, and $B = 1.5$ webers/meter^2.

The work required is the difference in energy between the two positions, or, from Eq. 33–12,

$$W = U_{\theta=180°} - U_{\theta=0} = (-\mu B \cos 180°) - (-\mu B \cos 0) = 2\mu B.$$

But $\mu = NiA$, so that

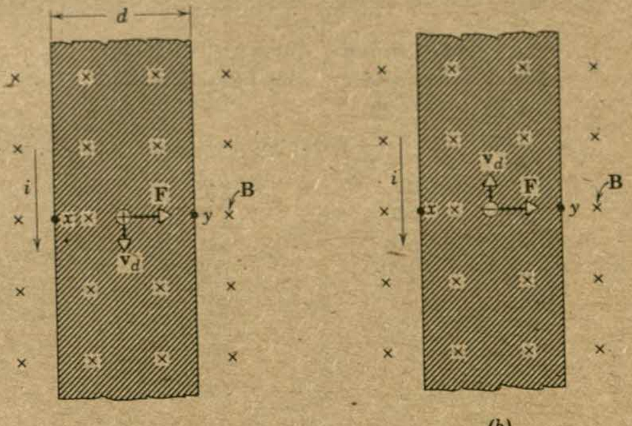
$$W = 2NiAB = 2Ni(\pi a^2)B$$

$$= (2)(100)(0.10 \text{ amp})(\pi)(5 \times 10^{-2} \text{ meter})^2(1.5 \text{ webers/meter}^2) = 0.24 \text{ joule} \blacktriangleleft$$

33-5 The Hall Effect

In 1879 E. H. Hall devised an experiment that gives the sign of the charge carriers in a conductor; see p. 771. Figure 33-10 shows a flat strip of copper, carrying a current i in the direction shown. As usual, the direction of the current arrow, labeled i, is the direction in which the charge carriers would move *if* they were positive. The current arrow can represent either positive charges moving down (as in Fig. 33-10*a*) or negative charges moving up (as in Fig. 33-10*b*). The Hall effect can be used to decide between these two possibilities.

A field of magnetic induction **B** is set up at right angles to the strip by placing the strip between the polefaces of an electromagnet. This field exerts a deflecting force **F** on the strip (given by $i\mathbf{l} \times \mathbf{B}$), which points to the right in the figure. Since the sideways force on the strip is due to the sideways forces on the charge carriers (given by $q\mathbf{v} \times \mathbf{B}$), it follows that these carriers, whether they are positive or negative, will tend to drift toward the right in Fig. 33-10 as they drift along the strip, producing a *transverse Hall potential difference* V_{xy} between points such as x and y. The sign of the charge carriers is determined by the sign of this Hall potential difference. If the carriers are positive, y will be at a higher potential than x; if they are negative, y will be at a lower potential than x. Experiment shows that in metals the charge carriers are negative.



(a) (b)

Fig. 33-10 A current i is set up in a copper strip placed in a field of magnetic induction **B**, assuming (*a*) positive carriers and (*b*) negative carriers.

To analyze the Hall effect quantitatively, let us use the *free-electron* model of a metal, the same model used in Section 31–4 to explain resistivity. The charge carriers can be assumed to move along the conductor with a certain constant drift speed v_d. The magnetic deflecting force that causes the moving charge carriers to drift toward the right edge of the strip is given by $q\mathbf{v}_d \times \mathbf{B}$ (see Eq. 33–3*a*).

The charge carriers do not build up without limit on the right edge of the strip because the displacement of charge gives rise to a transverse *Hall electric field* $\mathbf{E}_H$, which acts, inside the conductor, to oppose the sideways drift of the carriers. This Hall electric field is another manifestation of the Hall potential difference and is related to it by

$$E_H = V_{xy}/d.$$

Eventually an equilibrium is reached in which the sideways magnetic deflecting force on the charge carriers is just canceled by the oppositely directed electric force $q\mathbf{E}_H$ caused by the Hall electric field, or

$$q\mathbf{E}_H + q\mathbf{v}_d \times \mathbf{B} = 0,$$

which can be written

$$\mathbf{E}_H = -\mathbf{v}_d \times \mathbf{B}. \tag{33-13}$$

This equation shows explicitly that if $\mathbf{E}_H$ and $\mathbf{B}$ are measured $\mathbf{v}_d$ can be determined both in magnitude and direction; given the direction of $\mathbf{v}_d$, the sign of the charge carriers follows at once, as Fig. 33–10 shows.

The number of charge carriers per unit volume (n) can also be found from Hall effect measurements. If we write Eq. 33–13 in terms of magnitudes, for the case in which $\mathbf{v}_d$ and $\mathbf{B}$ are at right angles, we obtain $E_H = v_dB$. Combining this with Eq. 31–5 ($v_d = j/ne$) leads to

$$E_H = \frac{j}{ne}B \qquad \text{or} \qquad n = \frac{jB}{eE_H}. \tag{33-14}$$

The agreement between experiment and Eq. 33–14 is rather good for monovalent metals, as Table 33–1 shows.

Table 33–1

NUMBER OF CONDUCTION ELECTRONS PER UNIT VOLUME

Metal	Based on Hall Effect Data, $10^{22}/\text{cm}^3$	Calculated, Assuming One Electron/Atom, $10^{22}/\text{cm}^3$
Li	3.7	4.8
Na	2.5	2.6
K	1.5	1.3
Cs	0.80	0.85
Cu	11	8.4
Ag	7.4	6.0
Au	8.7	5.9

For nonmonovalent metals, for iron and similar magnetic materials, and for so-called semiconductors such as germanium, the simple interpretation of the Hall effect in terms of the free-electron model is not valid. A theoretical interpretation of the

Hall effect based on modern quantum physics gives a reasonable agreement with experiment in all cases.

▶ **Example 5.** A copper strip 2.0 cm wide and 1.0 mm thick is placed in a magnetic field with $B = 1.5$ webers/meter^2, as in Fig. 33-10. If a current of 200 amp is set up in the strip, what Hall potential difference appears across the strip?

From Eq. 33-14,

$$E_H = \frac{jB}{ne},$$

but

$$E_H = \frac{V_{xy}}{d} \quad \text{and} \quad j = \frac{i}{A} = \frac{i}{dh},$$

where h is the thickness of the strip. Combining these equations gives

$$V_{xy} = \frac{iB}{neh} = \frac{(200\text{ amp})(1.5\text{ webers/meter}^2)}{(8.4 \times 10^{28}/\text{meter}^3)(1.6 \times 10^{-19}\text{ coul})(1.0 \times 10^{-3}\text{ meter})}$$

$$= 2.2 \times 10^{-5}\text{ volt} = 22\mu\text{v}.$$

These potential differences are not large. See p. 681 for the calculation of n. ◀

33-6 Circulating Charges

Figure 33-11 shows a negatively charged particle introduced with velocity **v** into a uniform field of magnetic induction **B**. We assume that **v** is at right angles to **B** and thus lies entirely in the plane of the figure. The relation $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$ (Eq. 33-3*a*) shows that the particle will experience a sideways deflecting force of magnitude qvB. This force will lie in the plane of the figure, which means that the particle cannot leave this plane.

This reminds us of a stone held by a rope and whirled in a horizontal circle on a smooth surface. Here, too, a force of constant magnitude, the tension in the rope, acts in a plane and at right angles to the velocity. The charged particle, like the stone, also moves with constant speed in a circular

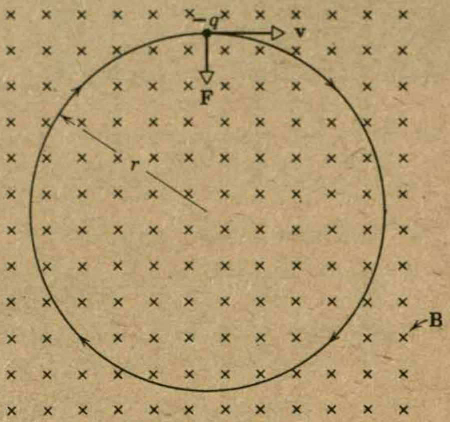


Fig. 33-11 A charge $-q$ circulates at right angles to a uniform magnetic field.

path. From Newton's second law we have

$$qvB = \frac{mv^2}{r} \quad \text{or} \quad r = \frac{mv}{qB}, \tag{33-15}$$

which gives the radius of the path. The three spirals in Fig. 33-4 show relatively low-energy electrons in a bubble chamber. The paths are not circles because the electrons lose energy by collisions in the chamber as they move.

The angular velocity ω is given by v/r or, from Eq. 33-15,

$$\omega = \frac{v}{r} = \frac{qB}{m}.$$

The frequency ν measured, say, in rev/sec, is given by

$$\nu = \frac{\omega}{2\pi} = \frac{qB}{2\pi m}. \tag{33-16}$$

Note that ν *does not depend on the speed of the particle.* Fast particles move in large circles (Eq. 33-15) and slow ones in small circles, but all require the same time T (the *period*) to complete one revolution in the field.

The frequency ν is a characteristic frequency for the charged particle in the field and may be compared to the characteristic frequency of a swinging pendulum in the earth's gravitational field or to the characteristic frequency of an oscillating mass-spring system. It is sometimes called the *cyclotron frequency* of the particle in the field because particles circulate at this frequency in the cyclotron.

▶ **Example 6.** A 10-ev electron is circulating in a plane at right angles to a uniform field of magnetic induction of 1.0×10^{-4} weber/meter² (= 1.0 gauss).

(*a*) What is its orbit radius?

The velocity of an electron whose kinetic energy is K can be found from

$$v = \sqrt{\frac{2K}{m}}.$$

The student should verify that this yields 1.9×10^6 meters/sec for v. Then, from Eq. 33-15,

$$r = \frac{mv}{qB} = \frac{(9.1 \times 10^{-31}\ \text{kg})(1.9 \times 10^6\ \text{meters/sec})}{(1.6 \times 10^{-19}\ \text{coul})(1.0 \times 10^{-4}\ \text{weber/meter}^2)} = 0.11\ \text{meter} = 11\ \text{cm}.$$

(*b*) What is the cyclotron frequency? From Eq. 33-16,

$$\nu = \frac{qB}{2\pi m} = \frac{(1.6 \times 10^{-19}\ \text{coul})(1.0 \times 10^{-4}\ \text{weber/meter}^2)}{(2\pi)(9.1 \times 10^{-31}\ \text{kg})} = 2.8 \times 10^6\ \text{rev/sec}.$$

(*c*) What is the period of revolution T?

$$T = \frac{1}{\nu} = \frac{1}{2.8 \times 10^6\ \text{rev/sec}} = 3.6 \times 10^{-7}\ \text{sec}.$$

Thus an electron requires 0.36 μsec to make 1 revolution in a 1.0-gauss field.

(*d*) What is the direction of circulation as viewed by an observer sighting along the field?

In Fig. 33-11 the magnetic force $q\mathbf{v} \times \mathbf{B}$ must point radially inward, since it provides the centripetal force. Since **B** points into the plane of the paper, **v** would have to point to the left at the position shown in the figure if the charge q were positive. However, the charge is an electron, with $q = -e$, which means that **v** must point to the right. Thus the charge circulates clockwise as viewed by an observer sighting in the direction of **B**. ◀

33-7 The Cyclotron

The cyclotron, first put into operation by Ernest Lawrence (1902–1958) in 1932, accelerates charged particles, such as protons or deuterons,* to high energies so that they can be used in atom-smashing experiments. Figure 33-12 shows the University of Pittsburgh cyclotron.

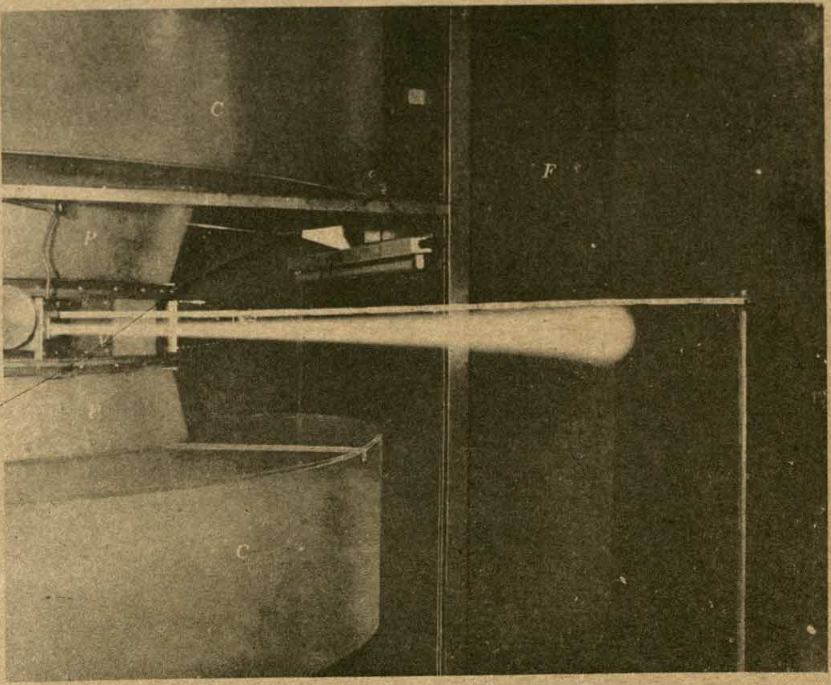


Fig. 33-12 The University of Pittsburgh cyclotron. Note vacuum chamber V, magnet frame F, magnetic pole faces P, magnet coils C, and the deuteron beam emerging into the air of the laboratory. The rule is 6 ft long. (Courtesy A. J. Allen.)

In an *ion source* at the center of the cyclotron molecules of deuterium are bombarded with electrons whose energy is high enough (say 100 ev) so that many positive ions are formed during the collisions. Many of these ions are free deuterons, which enter the cyclotron proper through a small hole in the wall of the ion source and are available to be accelerated.

---

* Deuterons are the nuclei of heavy hydrogen.

The cyclotron uses a modest potential difference for accelerating (say 10^5 volts), but it requires the ion to pass through this potential difference a number of times. To reach 10 Mev with 10^5 volts accelerating potential requires 100 passages. A magnetic field is used to bend the ions around so that they may pass again and again through the same accelerating potential.

Figure 33–13 is a top view of the part of the cyclotron that is inside the vacuum tank marked V in Fig. 33–12. The two D-shaped objects, called *dees*, are made of copper sheet and form part of an electric oscillator which establishes an accelerating potential difference across the gap between the dees. The direction of this potential difference is made to change sign some millions of times per second.

The dees are immersed in a magnetic field ($B \cong 1.6$ webers/meter^2) whose direction is out of the plane of Fig. 33–13. The field is set up by a large electromagnet, marked F in Fig. 33–12. Finally, the space in which the ions move is evacuated to a pressure of about 10^{-6} mm-Hg. If this were not done, the ions would continually collide with air molecules.

Suppose that a deuteron, emerging from the ion source, finds the dee that it is facing to be negative; it will accelerate toward this dee and will enter it. Once inside, it is screened from electrical forces by the metal walls of the dees. The magnetic field is not screened by the dees so that the ion bends in a circular path whose radius, which depends on the velocity, is given by Eq. 33–15, or

$$r = \frac{mv}{qB}.$$

After a time t_0 the ion emerges from the dee on the other side of the ion source. Let us assume that the accelerating potential has now changed sign. Thus the ion *again* faces a negative dee, is further accelerated, and again describes a semicircle, of somewhat larger radius (see Eq. 33–15), in the dee.

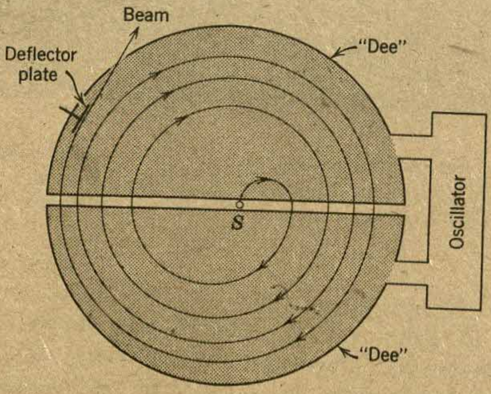


Fig. 33–13 The elements of a cyclotron showing the ion source S and the dees. The deflector plate, held at a suitable negative potential, deflects the particles out of the dee system.

The time of passage through this dee, however, is still t_0. This follows because the period of revolution T of an ion circulating in a magnetic field does not depend on the speed of the ion; see Eq. 33-16. This process goes on until the ion reaches the outer edge of one dee where it is pulled out of the system by a negatively charged deflector plate.

The key to the operation of the cyclotron is that the characteristic frequency ν at which the ion circulates in the field must be equal to the fixed frequency ν_0 of the electric oscillator. or

$$\nu = \nu_0.$$

This *resonance condition* says that if the energy of the circulating ion is to increase energy must be fed to it at a frequency ν_0 that is equal to the natural frequency ν at which the ion circulates in the field. In the same way we feed energy to a swing by pushing it at a frequency equal to the natural frequency of oscillation of the swing.

From Eq. 33-16 ($\nu = qB/2\pi m$), we can rewrite the resonance equation as

$$\frac{qB}{2\pi m} = \nu_0. \tag{33-17}$$

Once we have selected an ion to be accelerated, q/m is fixed; usually the oscillator is designed to work at a single frequency ν_0. We then "tune" the cyclotron by varying B until Eq. 33-17 is satisfied and an accelerated beam appears.

The *energy* of the particles produced in the cyclotron depends on the radius R of the dees. From Eq. 33-15 ($r = mv/qB$), the velocity of a particle circulating at this radius is given by

$$v = \frac{qBR}{m}.$$

The kinetic energy is then

$$K = \tfrac{1}{2}mv^2 = \frac{q^2B^2R^2}{2m}. \tag{33-18}$$

▶ **Example 7.** The University of Pittsburgh cyclotron has an oscillator frequency of 12×10^6 cycles/sec and a dee radius of 21 in. (*a*) What value of magnetic induction B is needed to accelerate deuterons?

From Eq. 33-17, $\nu_0 = qB/2\pi m$, so that

$$B = \frac{2\pi\nu_0 m}{q} = \frac{(2\pi)(12 \times 10^6/\text{sec})(3.3 \times 10^{-27}\ \text{kg})}{1.6 \times 10^{-19}\ \text{coul}} = 1.6\ \text{webers/meter}^2.$$

Note that the deuteron has the same charge as the proton but (very closely) twice the mass.

(*b*) What deuteron energy results?

From Eq. 33-18,

$$K = \frac{q^2B^2R^2}{2m} = \frac{(1.6 \times 10^{-19}\ \text{coul})^2(1.6\ \text{webers/meter}^2)^2(21 \times 0.0254\ \text{meter})^2}{(2)(3.3 \times 10^{-27}\ \text{kg})}$$

$$= (2.8 \times 10^{-12}\ \text{joule})\left(\frac{1\ \text{ev}}{1.6 \times 10^{-19}\ \text{joule}}\right) = 17\ \text{Mev}.$$ ◀

The cyclotron fails to operate at high energies because one of its assumptions, that the frequency of rotation of an ion circulating in a magnetic field is independent of its

speed, is true only for speeds much less than that of light. As the particle speed increases, we must use the *relativistic mass* m in Eq. 33–16. The relativistic mass increases with velocity (Eq. 8–20) so that at high enough speeds ν decreases with velocity. Thus the ions get out of step with the electric oscillator, and eventually the energy of the circulating ion stops increasing.

Another difficulty associated with the acceleration of particles to high energies is that the size of the magnet that would be required to guide such particles in a circular path is very large. For a 30-Bev proton, for example, in a field of 15,000 gauss the radius of curvature is 65 meters. A magnet of the cyclotron type of this size (about 430 ft in diameter) would be prohibitively expensive. Incidentally, a 30-Bev proton has a speed equal to 0.99998 that of light.

Both the relativistic and the economic limitations have been removed by techniques that can be understood in terms of Eq. 33–17 in which m is now taken to be the relativistic mass, given by Eq. 8–13, or

$$m = \frac{m_0}{\sqrt{1 - (v/c)^2}},$$

v being the speed of the particle and c being that of light.

As the particle speed increases, the relativistic mass m also increases. To maintain the equality in Eq. 33–17, and thus insure resonance, one may decrease the oscillator frequency ν_0 as the particle (assumed to be a proton) accelerates in such a way that the product $\nu_0 m$ remains constant. Accelerators that use this technique are called *synchrocyclotrons.*

To remove the magnet cost limitation one can vary *both* B and ν_0 in a cyclic fashion in such a way that not only is Eq. 33–17 satisfied at all times but the particle orbit radius remains constant during the acceleration process. This permits the use of an annular (or ring-shaped) magnet, rather than the conventional cyclotron type, at great saving in cost. With the *two* variables B and ν_0 at our disposal, it is possible to preserve *two* equalities during the acceleration process, one being Eq. 33–17 and the other being the relation

$$v = \omega_0 R_0 = (2\pi\nu_0)R_0$$

in which R_0 is the desired (fixed) orbit radius. Accelerators that use this technique are called *synchrotrons.* Table 33–2 shows some characteristics of the accel-

Table 33–2

THE CERN PROTON SYNCHROTRON

Orbit diameter	560 ft
Vacuum chamber cross section	5.5 in. × 2.7 in.
Maximum magnetic field	1.4 webers/meter^2
Frequency range per cycle	7 mc/sec
Pulse rate	20/min
Maximum proton energy	28 Bev
Energy gain per cycle	54 kev
Distance traveled by a proton	5×10^4 miles
Protons per pulse	10^{11}
Cost	\$28 × 10^6

erator built at Geneva, Switzerland, by the European Council for Nuclear Research (CERN), and embodying these principles.

33-8 Thomson's Experiment

In 1897 J. J. Thomson, working at the Cavendish Laboratory in Cambridge, measured the ratio of the charge e of the electron to its mass m by observing its deflection in combined electric and magnetic fields. The discovery of the electron is usually said to date from this historic experiment, although H. A. Lorentz and P. Zeeman (1865–1943), during the previous year, had measured this same quantity for electrons bound in atoms, using a method entirely different from Thomson's.

In Fig. 33-14, which is a modernized version of Thomson's apparatus, electrons are emitted from hot filament F and accelerated by an applied potential difference V. They then enter a region in which they move at right angles to an electric field **E** and a field of magnetic induction **B**; **E** and **B** are themselves at right angles to each other. The beam is made visible as a spot of light when it strikes fluorescent screen S. The entire region in which the electrons move is highly evacuated so that collisions with air molecules will not occur.

The resultant force on a charged particle moving through an electric and a magnetic field is given by Eq. 33-4, or

$$\mathbf{F} = q_0\mathbf{E} + q_0\mathbf{v} \times \mathbf{B}.$$

Study of Fig. 33-14 shows that the electric field deflects the particle upward and the magnetic field deflects it downward. If these deflecting forces are to

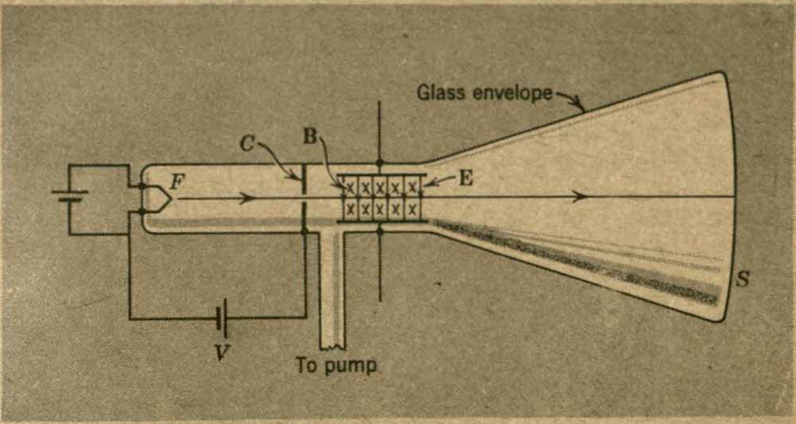


Fig. 33-14 Electrons from the heated filament F are accelerated by a potential difference V and pass through a hole in the screen C. After passing through a region in which perpendicular electric and magnetic fields are present, they strike the fluorescent screen S.

cancel (that is, if $\mathbf{F} = 0$), this equation, for this problem, reduces to

$$eE = evB$$

or

$$E = vB. \tag{33-19}$$

Thus for a given electron speed v the condition for zero deflection can be satisfied by adjusting E or B.

Thomson's procedure was (*a*) note the position of the undeflected beam spot, with **E** and **B** both equal to zero; (*b*) apply a fixed electric field **E**, measuring on the fluorescent screen the deflection so caused; and (*c*) apply a magnetic field and adjust its value until the beam deflection is restored to zero.

In Section 27–5 we saw that the deflection y of an electron in a purely electric field (step *b*), measured at the far edge of the deflecting plates, is given by Eq. 27–9, or, with small changes in notation,

$$y = \frac{eEl^2}{2mv^2},$$

where v is the electron speed and l is the length of the deflecting plates; y is not measurable directly, but it may be calculated from the measured displacement of the spot on the screen if the geometry of the apparatus is known. Thus y, E, and l are known; the ratio e/m and the velocity v are unknown. We cannot calculate e/m until we have found the velocity, which is the purpose of step *c* above.

If (step *c*) the electric force is set equal and opposite to the magnetic force, the net force is zero and we can write (Eq. 33–19)

$$v = \frac{E}{B}.$$

Substituting this equation into the equation for y and solving for the ratio e/m leads to

$$\frac{e}{m} = \frac{2yE}{B^2l^2}, \tag{33-20}$$

in which all the quantities on the right can be measured. Thomson's value for e/m was 1.7×10^{11} coul/kg, in excellent agreement with the modern value of 1.75890×10^{11} coul/kg.

QUESTIONS

1. Of the three vectors in the equation $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$, which pairs are always at right angles? Which may have any angle between them?

2. Why do we not simply define the magnetic induction **B** to point in the direction of the magnetic force that acts on the moving charge?

3. Imagine that you are sitting in a room with your back to one wall and that an electron beam, traveling horizontally from the back wall toward the front wall, is deflected to your right. What is the direction of the field of magnetic induction that exists in the room?

4. If an electron is not deflected in passing through a certain region of space, can we be sure that there is no magnetic field in that region?

5. If a moving electron is deflected sideways in passing through a certain region of space, can we be sure that a magnetic field exists in that region?

6. A beam of protons is deflected sideways. Could this deflection be caused (*a*) by an electric field? (*b*) By a magnetic field? (*c*) If either could be responsible, how would you be able to tell which was present?

7. A conductor, even though it is carrying a current, has zero net charge. Why, then, does a magnetic field exert a force on it?

8. Equation 33–11 ($\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B}$) shows that there is no torque on a current loop in an external magnetic field if the angle between the axis of the loop and the field is (*a*) 0° or (*b*) 180°. Discuss the nature of the equilibrium (that is, is it stable, neutral, or unstable?) for these two positions.

9. In Example 4 we showed that the work required to turn a current loop end-for-end in an external magnetic field is $2\mu B$. Does this hold no matter what the original orientation of the loop was?

10. Imagine that the room in which you are seated is filled with a uniform magnetic field with **B** pointing vertically upward. A circular loop of wire has its plane horizontal. For what direction of current in the loop, as viewed from above, will the loop be in stable equilibrium with respect to forces and torques of magnetic origin?

11. A rectangular current loop is in an arbitrary orientation in an external magnetic field. Is any work required to rotate the loop about an axis perpendicular to its plane?

12. (*a*) In measuring Hall potential differences, why must we be careful that points x and y in Fig. 33–10 are exactly opposite to each other? (*b*) If one of the contacts is movable, what procedure might we follow in adjusting it to make sure that the two points are properly located?

13. A uniform magnetic field fills a certain cubical region of space. Can an electron be fired into this cube from the outside in such a way that it will travel in a closed circular path inside the cube?

14. Imagine the room in which you are seated to be filled with a uniform magnetic field with **B** pointing vertically downward. At the center of the room two electrons are suddenly projected horizontally with the same speed but in opposite directions. (*a*) Discuss their motions. (*b*) Discuss their motions if one particle is an electron and one a positron.

15. In Fig. 33–4, why are the low-energy electron tracks spirals? That is, why does the radius of curvature change in the constant magnetic field in which the chamber is immersed?

16. What are the primary functions of (*a*) the electric field and (*b*) the magnetic field in the cyclotron?

17. For Thomson's e/m experiment to work properly (Section 33–8), is it essential that the electrons have a fairly constant speed?

PROBLEMS

1. The electrons in the beam of a television tube have an energy of 12 kev. The tube is oriented so that the electrons move horizontally from south to north. The vertical component of the earth's magnetic field points down and has $B = 5.5 \times 10^{-5}$ weber/meter^2. (*a*) In what direction will the beam deflect? (*b*) What is the acceleration of a given electron? (*c*) How far will the beam deflect in moving 20 cm through the television tube?

2. In a nuclear experiment a 1.0-Mev proton moves in a uniform magnetic field in a circular path. What energy must (*a*) an alpha particle and (*b*) a deuteron have if they are to circulate in the same orbit?

3. A wire 1.0 meter long carries a current of 10 amp and makes an angle of 30° with a uniform magnetic field with $B = 1.5$ webers/meter^2. Calculate the magnitude and direction of the force on the wire.

4. A wire of 60 cm length and mass 10 gm is suspended by a pair of flexible leads in a magnetic field of induction 0.40 weber/meter^2. What are the magnitude and direction of the current required to remove the tension in the supporting leads? See Fig. 33–15.

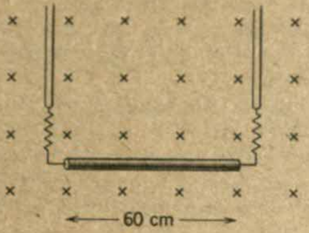


Fig. 33–15

5. Express magnetic induction B and magnetic flux Φ in terms of the fundamental dimensions M, L, T, and Q (mass, length, time, and charge).

6. A metal wire of mass m slides without friction on two rails spaced a distance d apart, as in Fig. 33–16. The track lies in a vertical uniform field of magnetic induction **B**. A *constant current* i flows from generator G along one rail, across the wire, and back down the other rail. Find the velocity (speed and direction) of the wire as a function of time, assuming it to be at rest at $t = 0$.

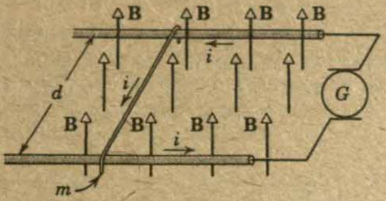


Fig. 33–16

7. A U-shaped wire of mass m and length l is immersed with its two ends in mercury (Fig. 33–17). The wire is in a homogeneous field of magnetic induction **B**. If a charge, that is, a current pulse $q = \int i\,dt$, is sent through the wire, the wire will jump up. Calculate, from the height h that the wire reaches, the size of the charge or current pulse, assuming that the time of the current pulse is very small in comparison with the time of flight. Make use of the fact that impulse of force equals $\int F\,dt$, which equals mv. (Hint: Try to relate $\int i\,dt$ to $\int F\,dt$.) Evaluate q for $B = 0.1$ weber/meter^2, $m = 10$ gm, $l = 20$ cm, and $h = 3$ meters.

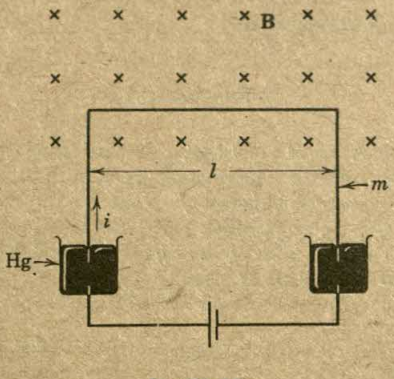


Fig. 33–17

8. Figure 33-18 shows a wire ring of radius a at right angles to the general direction of a radially symmetric diverging magnetic field. The magnetic induction at the ring is everywhere of the same magnitude B, and its direction at the ring is everywhere at an angle θ with a normal to the plane of the ring. The twisted lead wires have no effect on the problem. Find the magnitude and direction of the force the field exerts on the ring if the ring carries a current i as shown in the figure.

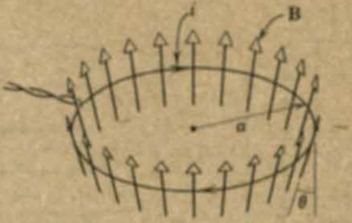


Fig. 33-18

9. A copper rod weighing 0.30 lb rests on two rails 1.0 ft apart and carries a current of 50 amp from one rail to the other. The coefficient of starting friction is 0.60. What is the smallest magnetic field that would cause the bar to slide and what is its direction?

10. Figure 33-19 shows a wire of arbitrary shape carrying a current i between points a and b. The wire lies in a plane at right angles to a uniform field of magnetic induction **B**. Prove that the force on the wire is the same as that on a straight wire carrying a current i directly from a to b. (Hint: Replace the wire by a series of "steps" parallel and perpendicular to the straight line joining a and b.)

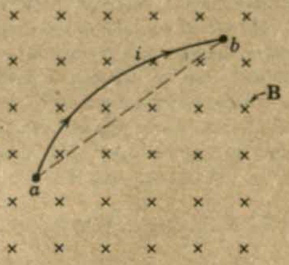


Fig. 33-19

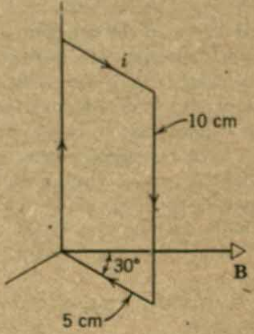


Fig. 33-20

11. Figure 33-20 shows a rectangular twenty-turn loop of wire, 10 cm by 5.0 cm. It carries a current of 0.10 amp and is hinged at one side. What torque (direction and magnitude) acts on the loop if it is mounted with its plane at an angle of 30° to the direction of a uniform field of magnetic induction 0.50 weber/meter^2?

12. Prove that the relation $\tau = NAiB \sin \theta$ holds for closed loops of arbitrary shape. (Hint: Replace the loop of arbitrary shape by an assembly of adjacent long, thin—approximately rectangular—loops which are equivalent to it as far as the distribution of current is concerned.)

13. A length L of wire carries a current i. Show that if the wire is formed into a circular coil the maximum torque in a given magnetic field is developed when the coil has *one* turn only and the maximum torque has the value

$$\tau = \frac{1}{4\pi} L^2 iB.$$

14. Figure 33–21 shows a wooden cylinder with a mass m of 0.25 kg, a radius R, and a length l of 0.1 meter with N equal to ten turns of wire wrapped around it longitudinally, so that the plane of the wire loop contains the axis of the cylinder. What is the least current through the loop that will prevent the cylinder from rolling down an inclined plane whose surface is inclined at an angle θ to the horizontal, in the presence of a vertical field of magnetic induction 0.5 weber/meter², if the plane of the windings is parallel to the inclined plane?

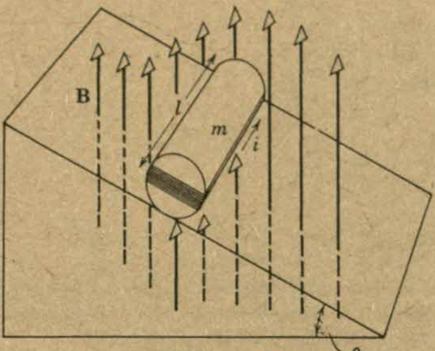


Fig. 33-21

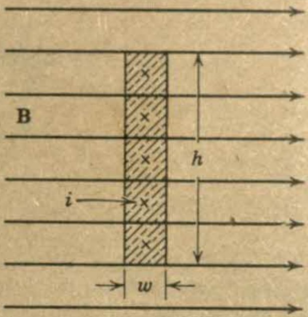


Fig. 33-22

15. A current i, indicated by the crosses in Fig. 33–22, is established in a strip of copper of height h and width w. A uniform field of magnetic induction **B** is applied at right angles to the strip. (*a*) Calculate the drift speed v_d for the electrons. (*b*) What are the magnitude and direction of the magnetic force **F** acting on the electrons? (*c*) What would the magnitude and direction of a homogeneous electric field **E** have to be in order to counterbalance the effect of the magnetic field? (*d*) What is the voltage V necessary between two sides of the conductor in order to create this field **E**? Between which sides of the conductor would this voltage have to be applied? (*e*) If no electric field is applied *from the outside*, the electrons will be pushed somewhat to one side and therefore will give rise to a uniform electric field $\mathbf{E}_H$ across the conductor until the forces of this electrostatic field $\mathbf{E}_H$ balance the magnetic forces encountered in part (*b*). What will be the magnitude and direction of the field $\mathbf{E}_H$? Assume that n, the number of conduction electrons per unit volume, is 1.1×10^{29}/meter³ and that $h = 0.02$ meter, $w = 0.1$ cm, $i = 50$ amp, and $B = 2$ webers/meter².

16. (*a*) Show that the ratio of the Hall electric field E_H to the electric field E responsible for the current is

$$\frac{E_H}{E} = \frac{B}{ne\rho}.$$

(*b*) What is the angle between $\mathbf{E}_H$ and **E**? (*c*) Evaluate this ratio for the conditions of Example 5.

17. A proton, a deuteron, and an α-particle, accelerated through the same potential difference, enter a region of uniform magnetic field, moving at right angles to **B**. (*a*) Compare their kinetic energies. (*b*) If the radius of the proton's circular path is 10 cm, what are the radii of the deuteron and the α-particle paths?

18. A proton, a deuteron, and an α-particle with the same kinetic energies enter a region of uniform magnetic field, moving at right angles to **B**. Compare the radii of their circular paths.

19. An α-particle travels in a circular path of radius 0.45 meter in a magnetic field with $B = 1.2$ webers/meter². Calculate (*a*) its speed, (*b*) its period of revolution, (*c*) its kinetic energy, and (*d*) the potential difference through which it would have to be accelerated to achieve this energy.

20. An electron is accelerated through 15,000 volts and is then allowed to circulate at right angles to a uniform magnetic field with $B = 250$ gauss. What is its path radius?

21. Electrons are observed to be ejected in various directions with negligible speed from the negative plate of a parallel-plate capacitor when the plate is illuminated by light of a certain wavelength (photoelectric effect). The plates are separated by a distance d and a potential difference V is maintained between them. Show that none of these electrons will reach the positive plate if a magnetic field is applied at right angles to the electric field and the magnetic induction has a value

$$B > \left(\frac{2Vm}{ed^2}\right)^{1/2},$$

in which m and e are the electron mass and charge, respectively.

22. Show that the radius of curvature of a charged particle moving at right angles to a magnetic field is proportional to its momentum.

23. What is the smallest magnetic field (magnitude and direction) that can be set up at the equator to permit a proton of speed 1.0×10^7 meters/sec to circulate around the earth?

24. A deuteron in a large cyclotron is moving in a magnetic field with $B = 1.5$ webers/meter^2 and an orbit radius of 2.0 meters. Because of a grazing collision with a target, the deuteron breaks up, with a negligible loss of kinetic energy, into a proton and a neutron. Discuss the subsequent motions of each. Assume that the deuteron energy is shared equally by the proton and neutron at breakup.

25. A 2-kev positron is projected into a uniform field of induction **B** of 0.10 weber/meter^2 with its velocity vector making an angle of 89° with **B**. Convince yourself that the path will be a helix, its axis being the direction of **B**. Find the period, the pitch p, and the radius r of the helix; see Fig. 33-23.

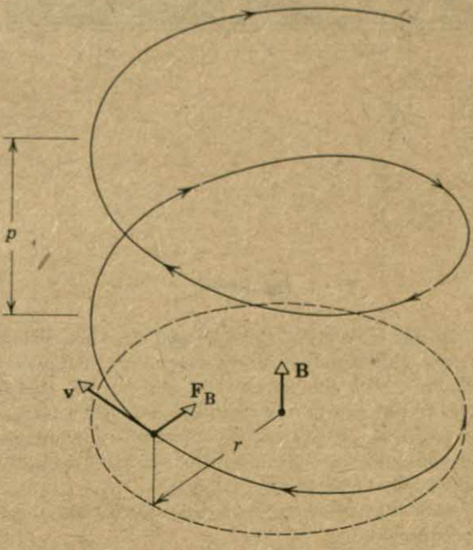


Fig. 33-23

26. (*a*) In a magnetic field with $B = 0.50$ weber meter^2, for what path radius will an electron circulate at 0.1 the speed of light? (*b*) What will its kinetic energy be?

27. *Time-of-flight spectrometer.* S. A. Goudsmit has devised a method for measuring accurately the masses of heavy ions by timing their period of circulation in a known magnetic field. A singly charged ion of iodine makes 7 rev in a field of 4.5×10^{-2} weber/meter^2 in about 1.29×10^{-3} sec. What (approximately) is its mass in kilograms? Actually, the mass measurements are carried out to much greater accuracy than these approximate data suggest.

28. *Mass spectrometer.* Figure 33–24 shows an arrangement used by Dempster to measure the masses of ions. An ion of mass M and charge $+q$ is produced essentially at rest in source S, a chamber in which a gas discharge is taking place. The ion is accelerated by potential difference V and allowed to enter a field of magnetic induction $\mathbf{B}$. In the field it moves in a semicircle, striking a photographic plate at distance x from the entry slit and being recorded. Show that the mass M is given by

$$M = \frac{B^2 q}{8V} x^2.$$

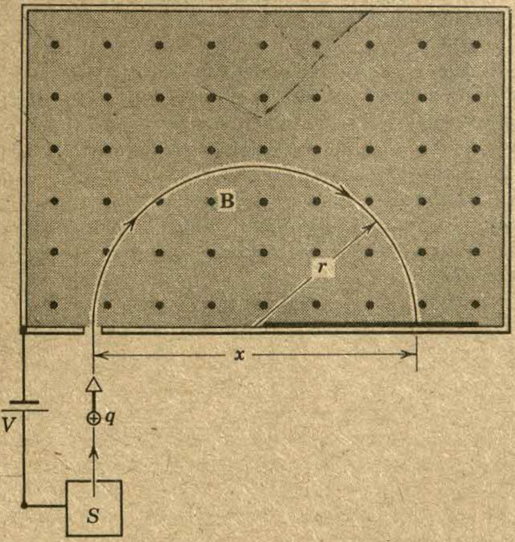


Fig. 33–24

29. *Zeeman effect.* In Bohr's theory of the hydrogen atom the electron can be thought of as moving in a circular orbit of radius r about the proton. Suppose that such an atom is placed in a magnetic field, with the plane of the orbit at right angles to $\mathbf{B}$. (*a*) If the electron is circulating clockwise, as viewed by an observer sighting along $\mathbf{B}$, will the angular frequency increase or decrease? (*b*) What if the electron is circulating counterclockwise? Assume that the orbit radius does not change. [Hint: The centripetal force is now partially electric ($\mathbf{F}_E$) and partially magnetic ($\mathbf{F}_B$) in origin.]

30. In Problem 29 show that the change in frequency of rotation caused by the magnetic field is given approximately by

$$\Delta\nu = \pm \frac{Be}{4\pi m}.$$

Such frequency shifts were actually observed by Zeeman in 1896. (Hint: Calculate the frequency of rotation without the magnetic field and also with it. Subtract, bearing in mind that because the effect of the magnetic field is very small some—but not all—terms containing B can be set equal to zero with little error.)

31. The University of Pittsburgh cyclotron is normally adjusted to accelerate deuterons. (*a*) What energy of protons could it produce, using the same oscillator frequency as that used for deuterons? (*b*) What magnetic induction would be required? (*c*) What energy of protons could be produced if the magnetic induction was left at the value used for deuterons? (*d*) What oscillator frequency would then be required? (*e*) Answer the same questions for α-particles.

32. Estimate the total path length traversed by a deuteron in the University of Pittsburgh cyclotron during the acceleration process. Assume an accelerating potential between the dees of 80,000 volts.

33. In a synchrocyclotron producing 400-Mev protons, what must the ratio be of the oscillator frequency at the beginning of an accelerating cycle to that at the end? Such a proton has a speed of $0.70c$, where c is the speed of light.

34. A 10-kev electron moving horizontally enters a region of space in which there is a downward-directed electric field of magnitude 100 volts/cm. (*a*) What are the magnitude and direction of the (smallest) field of magnetic induction that will allow the electron to continue to move horizontally? Ignore gravitational forces, which are rather small. (*b*) Is it possible for a proton to pass through this combination of fields undeflected? If so, under what circumstances?

35. An electric field of 1500 volts/meter and a magnetic field of 0.40 weber/meter^2 act on a moving electron to produce no force. (*a*) Calculate the minimum electron speed v. (*b*) Draw the vectors $\mathbf{E}$, $\mathbf{B}$, and $\mathbf{v}$.

Ampère's Law

CHAPTER 34

34-1 Ampère's Law

One class of problems involving magnetic fields, dealt with in Chapter 33, concerns the forces exerted by a magnetic field on a moving charge or on a current-carrying conductor and the torque exerted on a magnetic dipole. A second class concerns the *production* of a magnetic field by a current-carrying conductor or by moving charges. This chapter deals with problems of this second class.

The discovery that currents produce magnetic effects was made by Oersted in 1820. Figure 34-1, which shows a wire surrounded by a number of small magnets, illustrates a modification of his experiment. If there is no current in the wire, all the magnets are aligned with the earth's magnetic field. When a strong current is present, the magnets point so as to suggest that the magnetic lines of induction form closed circles around the wire. This view is strengthened by the experiment of Fig. 34-2, which shows iron filings on a horizontal glass plate, through the center of which a current-carrying conductor passes.

Today we write the quantitative relationship between current i and the magnetic field **B** as

$$\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0 i, \tag{34-1}$$

which is known as *Ampère's law*. Ampère, being an advocate of the action-at-a-distance point of view, did not formulate his results in terms of fields; this was first done by Maxwell. Ampère's law, including an important ex-

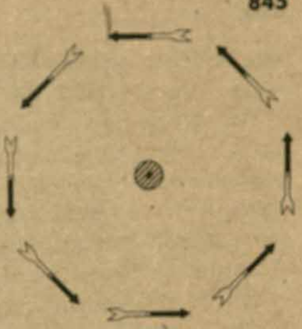


Fig. 34-1 An array of compass needles near a wire carrying a strong current. The black ends of the compass needles are their north poles. The dot shows the current emerging from the page. As usual, the direction of a current is taken as the direction of flow of positive charge.

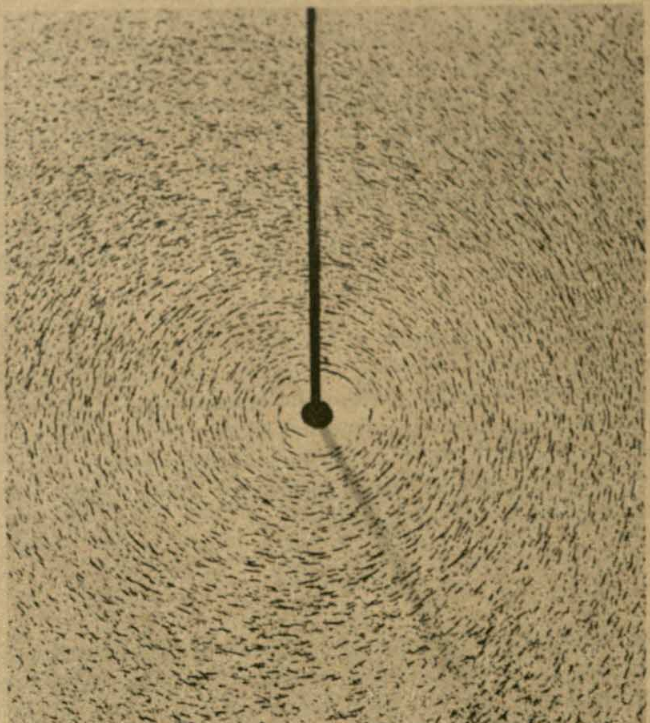


Fig. 34-2 Iron filings around a wire carrying a strong current. (Courtesy Physical Science Study Committee.)

tension of it made later by Maxwell, is one of the basic equations of electromagnetism (see Table 38–3).

We can gain an appreciation of the way Ampère's law developed historically by considering a hypothetical experiment which has, in fact, much in common with experiments that were actually carried out. The experiment consists of measuring **B** at various distances r from a long straight wire of circular cross section and carrying a current i. This can be done by making quantitative the qualitative observation of Fig. 34–1.

Let us put a small compass needle a distance r from the wire. Such a needle, a small magnetic dipole, tends to line up with an external magnetic field, with its north pole pointing in the direction of **B**. Figure 34–1 makes it clear that **B** at the site of the dipole is tangent to a circle of radius r centered on the wire.

If the current in the wire of Fig. 34–1 is reversed in direction, all the compass needles would reverse end-for-end. This experimental result leads to the "right-hand rule" for finding the direction of **B** near a wire carrying a current i: *Grasp the wire with the right hand, the thumb pointing in the direction of the current. The fingers will curl around the wire in the direction of* **B**.

Let us now turn the dipole through an angle θ from its equilibrium position. To do this, we must exert an external torque just large enough to overcome the restoring torque $\boldsymbol{\tau}$ that will act on the dipole. $\boldsymbol{\tau}$, θ, and **B** are related by Eq. 33–11 ($\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B}$), which can be written in terms of magnitudes as

$$\tau = \mu B \sin\theta \tag{34–2}$$

and in which μ is the magnitude of the magnetic moment of the dipole, θ being the angle between the vectors $\boldsymbol{\mu}$ and **B**. Even though we may not know the value of μ for the compass needle, we may take it to be a constant, independent of the position or orientation of the needle. Thus by measuring τ and θ in Eq. 34–2 we can obtain a *relative* measure of B for various distances r and for various currents i in the wire. The experimental results can be described by the proportionality

$$B \propto \frac{i}{r}. \tag{34–3}$$

We can convert this proportionality into an equality by inserting a proportionality constant. As for Coulomb's law, and for similar reasons (see Section 26–4), we do not write this constant simply as k but in a more complex form, namely $\mu_0/2\pi$, in which μ_0 is called the *permeability constant.** Equation 34–3 then becomes

$$B = \frac{\mu_0 i}{2\pi r}, \tag{34–4}$$

which we choose to write in the form

$$(B)(2\pi r) = \mu_0 i. \tag{34–5}$$

---

* μ_0 has no connection with the dipole moment μ that appears in Eq. 34–2.

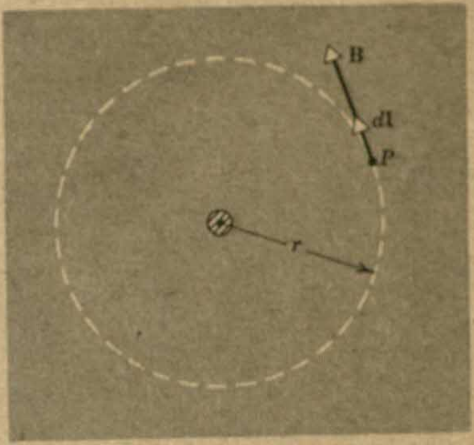


Fig. 34-3 A circular path of integration surrounding a wire. The central dot suggests a current i in the wire emerging from the page. Note that the angle between $\mathbf{B}$ and $d\mathbf{l}$ is zero so that $\mathbf{B}\cdot d\mathbf{l} = B\,dl$.

The left side of Eq. 34-5 can easily be shown to be $\oint \mathbf{B}\cdot d\mathbf{l}$ for a path consisting of a circle of radius r centered on the wire. For all points on this circle $\mathbf{B}$ has the same (constant) magnitude B and $d\mathbf{l}$, which is always tangent to the path of integration, points in the same direction as $\mathbf{B}$, as Fig. 34-3 shows. Thus

$$\oint \mathbf{B}\cdot d\mathbf{l} = \oint B\,dl = B\oint dl = (B)(2\pi r),$$

$\oint dl$ being simply the circumference of the circle. *In this special case,* therefore, we can write the experimentally observed connection between the field and the current as

$$\oint \mathbf{B}\cdot d\mathbf{l} = \mu_0 i, \tag{34-1}$$

which is Ampère's law. A host of other experiments suggests that Eq. 34-1 is true in general * for *any* magnetic field configuration, for *any* assembly of currents, and for *any* path of integration.

In applying Ampère's law in the general case, we construct a *closed linear path* in the magnetic field as shown in Fig. 34-4. This path is divided into elements of length $d\mathbf{l}$, and for each element the quantity $\mathbf{B}\cdot d\mathbf{l}$ is evaluated. Recall that $\mathbf{B}\cdot d\mathbf{l}$ has the magnitude $B\,dl\cos\theta$ and can be interpreted as the product of dl and the component of $\mathbf{B}$ $(= B\cos\theta)$ parallel to $d\mathbf{l}$. The inte-

---

* We must modify Eq. 34-1 if a time-varying electric field is present within the path of integration. In this chapter we assume that if electric fields are present, they are constant in magnitude and direction.

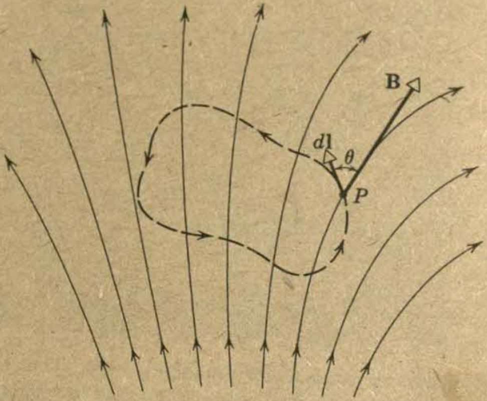


Fig. 34–4 A path of integration in a magnetic field.

gral is the sum of the quantities $\mathbf{B}\cdot d\mathbf{l}$ for all path elements in the complete loop. The term i on the right of Eq. 34–1 is the *net* current that passes through the area bounded by the closed path.

The permeability constant in Ampère's law has an assigned value of

$$\mu_0 = 4\pi \times 10^{-7} \text{ weber/amp-meter}$$

Both this and the permittivity constant (ϵ_0) occur in electromagnetic formulas when the mks system of units is used.

The student may wonder why ϵ_0 in Coulomb's law is a measured quantity, whereas μ_0 in Ampère's law is an assigned quantity. The answer is that the ampere, which is the mks unit for the current i in Ampère's law, is defined by a laboratory technique (the *current balance*) that involves forces exerted by magnetic fields and in which this same constant μ_0 appears. In effect, as we show in detail in Section 34–4, the size of the current that we agree to define as one ampere is adjusted so that μ_0 may have exactly the value assigned to it above. In Coulomb's law, on the other hand, the quantities $\mathbf{F}$, q, and r are measured in ways in which the constant ϵ_0 plays no role. This constant must then take on the particular value that makes the left side of Coulomb's law equal to the right side; no arbitrary assignment is possible.

34–2 B Near a Long Wire

We have seen that the lines of magnetic induction for a long straight wire carrying a current i are concentric circles centered on the wire and that B at a distance r from the wire is given by Eq. 34–4:

$$B = \frac{\mu_0 i}{2\pi r}. \tag{34–4}$$

We may regard this as an experimental result consistent with, and readily derivable from, Ampère's law.

It is interesting to compare Eq. 34-4 with the expression for the electric field near a long line of charge, or

$$E = \frac{1}{2\pi\epsilon_0}\frac{\lambda}{r}. \qquad (28\text{-}7)$$

In each case there are multiplying constants, namely $\mu_0/2\pi$ and $1/2\pi\epsilon_0$, and factors describing the device responsible for the field, namely i and λ. Finally, each field varies as $1/r$.

Equation 28-7 may be derived from Gauss's law by relating the electric field at a Gaussian surface to the net charge within this surface. The (surface) integral in Gauss's law is evaluated for a closed cylindrical surface to which the lines of **E** are everywhere perpendicular.

Equation 34-4 may be derived from Ampère's law by relating the magnetic field at a path of integration to the net current that pierces this path. The (line) integral in Ampère's law is evaluated for a closed circular path to which the lines of **B** are everywhere tangent.

▶ **Example 1** Derive an expression for **B** at a distance r from the center of a long cylindrical wire of radius R, where $r < R$. The wire carries a current i_0, distributed uniformly over the cross section of the wire.

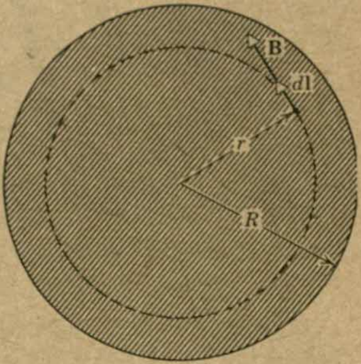


Fig. 34-5 Example 1. A circular path of integration inside a wire. A current i_0, distributed uniformly over the cross section of the wire, emerges from the page.

Figure 34-5 shows a circular path of integration inside the wire. Symmetry suggests that **B** is tangent to the path as shown. Ampère's law,

$$\oint \mathbf{B}\cdot d\mathbf{l} = \mu_0 i,$$

gives

$$(B)(2\pi r) = \mu_0 i_0 \frac{\pi r^2}{\pi R^2},$$

since only the fraction of the current that passes through the path of integration is included in the factor i on the right. Solving for B and dropping the subscript on the current yields

$$B = \frac{\mu_0 i r}{2\pi R^2}.$$

At the surface of the wire ($r = R$) this equation reduces to the same expression as that found by putting $r = R$ in Eq. 34-4 ($B = \mu_0 i/2\pi R$).

Example 2. Figure 34–6 shows a flat strip of copper of width a and negligible thickness carrying a current i. Find the magnetic field at a distance R from the center of the strip, at right angles to the strip.

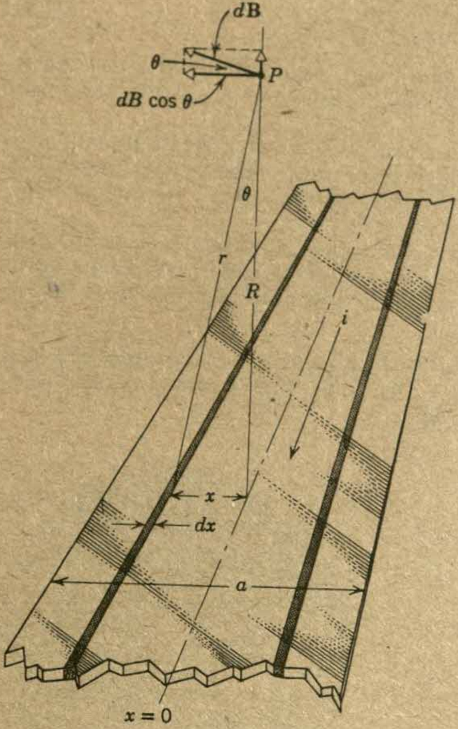


Fig. 34–6 Example 2. A flat strip of width a carries a current i.

Let us subdivide the strip into long infinitesimal filaments of width dx, each of which may be treated as a wire carrying a current di given by $i(dx/a)$. The field contribution dB at point P in Fig. 34–6 is given, for the element shown, by the differential form of Eq. 34–4, or

$$dB = \frac{\mu_0}{2\pi}\frac{di}{r} = \frac{\mu_0}{2\pi}\frac{i(dx/a)}{R \sec\theta}.$$

Note that the vector $d\mathbf{B}$ is at right angles to the line marked r.

Only the horizontal component of $d\mathbf{B}$, namely $dB \cos\theta$, is effective, the vertical component being canceled by the contribution associated with a symmetrically located filament on the other side of the origin. Thus B at point P is given by the (scalar) integral

$$B = \int dB \cos\theta = \int \frac{\mu_0 i(dx/a)}{2\pi R \sec\theta}\cos\theta$$

$$= \frac{\mu_0 i}{2\pi a R}\int \frac{dx}{\sec^2\theta}.$$

The variables x and θ are not independent, being related by

$$x = R \tan \theta$$

or

$$dx = R \sec^2 \theta \, d\theta.$$

Bearing in mind that the limits on θ are $\pm \tan^{-1} (a/2R)$ and eliminating dx from this expression for B, we find

$$B = \frac{\mu_0 i}{2\pi a R} \int \frac{R \sec^2 \theta \, d\theta}{\sec^2 \theta}$$

$$= \frac{\mu_0 i}{2\pi a} \int_{-\tan^{-1} a/2R}^{+\tan^{-1} a/2R} d\theta = \frac{\mu_0 i}{\pi a} \tan^{-1} \frac{a}{2R}.$$

At points far from the strip, $a/2R$ is a small angle, for which $\tan^{-1} \alpha \cong \alpha$. Thus we have, as an approximate result,

$$B \cong \frac{\mu_0 i}{\pi a} \left(\frac{a}{2R} \right) = \frac{\mu_0}{2\pi} \frac{i}{R}.$$

This result is expected because at distant points the strip cannot be distinguished from a cylindrical wire (see Eq. 34-4). ◀

34-3 Magnetic Lines of Induction

Figure 34-7 shows the lines of magnetic induction representing the field of **B** near a long straight wire. Note the increase in the spacing of the lines with increasing distance from the wire. This represents the $1/r$ decrease in B predicted by Eq. 34-4.

Figure 34-8 shows the resultant lines of magnetic induction associated with a current in a wire that is oriented at right angles to a uniform *external* field $\mathbf{B}_e$. At any point the resultant magnetic induction **B** will be the vector sum of $\mathbf{B}_e$ and $\mathbf{B}_i$, where $\mathbf{B}_i$ is the magnetic induction set up by the current in the wire. The fields $\mathbf{B}_e$ and $\mathbf{B}_i$ tend to cancel above the wire and to reenforce each other below the wire. At point P in Fig. 34-8 $\mathbf{B}_e$ and $\mathbf{B}_i$ cancel exactly. Very near the wire the field is represented by circular lines and is essentially $\mathbf{B}_i$.

Michael Faraday, who originated the concept of lines of induction, endowed them with more reality than they are currently given. He imagined

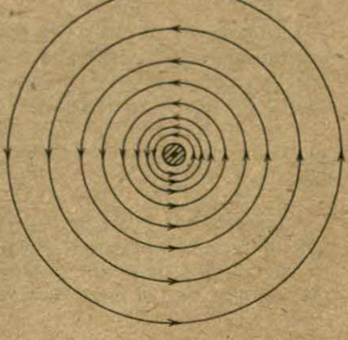


Fig. 34-7 Lines of **B** near a long cylindrical wire. A current i, suggested by the central dot, emerges from the page.

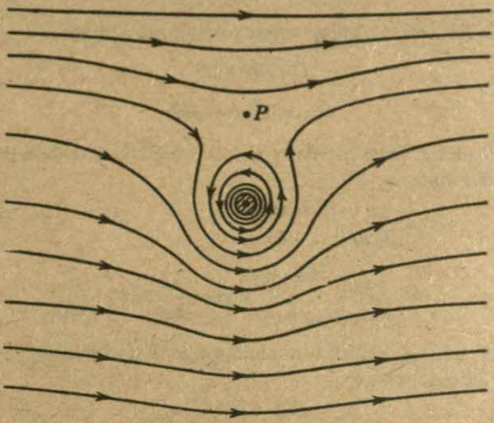


Fig. 34–8 Lines of **B** near a long current-carrying wire immersed in a uniform external field $\mathbf{B}_e$ that points to the right. The current i is emerging from the page.

that, like stretched rubber bands, they represent the site of mechanical forces. On this picture can we not visualize that the wire in Fig. 34–8 will be pushed up? Today we use lines of induction largely for purposes of visualization. For quantitative calculations we use the field vectors, describing the force on the wire in Fig. 34–8, for example, from the relation $\mathbf{F} = i\mathbf{l} \times \mathbf{B}$.

In applying this relation to Fig. 34–8, we recall that **B** is always the *external field* in which the wire is immersed; that is, it is $\mathbf{B}_e$ and thus points to the right. Since **l** points out of the page, the magnetic force on the wire ($= i\mathbf{l} \times \mathbf{B}_e$) does indeed point up. It is necessary to use only the external field in such calculations because the field set up by the current in the wire cannot exert a force on the wire, just as the gravitational field of the earth cannot exert a force on the earth itself but only on another body. In Fig. 34–7, for example, there is no magnetic force on the wire because no external magnetic field is present.

Faraday's idea of lines of induction was instrumental in overthrowing the older action-at-a-distance theory of magnetic (and electric) attraction. Like many new ideas, it was not immediately accepted. In 1851, for example, Faraday wrote:

> I cannot refrain from again expressing my conviction of the truthfulness of the representation, which the idea of lines of force affords in regard to magnetic action. All the points that are experimentally established in regard to that action—i.e., all that is not hypothetical—appear to be well and truly represented by it.

On the other hand, four years later another well-known British scientist, Sir George Airy, wrote:

> I declare that I can hardly imagine anyone who practically and numerically knows this agreement [with the action-at-a-distance theory] to hesitate an instant in the choice between this simple and precise action, on the one hand, and anything so vague as lines of force, on the other hand.

34-4 Two Parallel Conductors

Figure 34-9 shows two long parallel wires separated by a distance d and carrying currents i_a and i_b. It is an experimental fact, noted by Ampère only one week after word of Oersted's experiments reached Paris, that two such conductors attract each other.

Some of Ampère's colleagues thought that in view of Oersted's experiment this attraction between two conductors was an obvious result and did not need to be proved. They reasoned that if wire a and wire b each exert forces on a compass needle they should exert forces on each other. This conclusion is wrong. When he heard it, Arago, a contemporary of Ampère, drew two iron keys from his pocket and replied, "Each of these keys attracts a magnet. Do you believe that they therefore also attract each other?"

Wire a in Fig. 34-9 will produce a field of induction $\mathbf{B}_a$ at all nearby points. The magnitude of $\mathbf{B}_a$, due to the current i_a, at the site of the second wire is, from Eq. 34-4,

$$B_a = \frac{\mu_0 i_a}{2\pi d}.$$

The right-hand rule shows that the direction of $\mathbf{B}_a$ at wire b is down, as shown in the figure.

Wire b, which carries a current i_b, finds itself immersed in an *external* field of magnetic induction $\mathbf{B}_a$. A length l of this wire will experience a sideways magnetic force ($= i\mathbf{l} \times \mathbf{B}$) whose magnitude is

$$F_b = i_b l B_a = \frac{\mu_0 l i_b i_a}{2\pi d}. \tag{34-6}$$

The vector rule of signs tells us that $\mathbf{F}_b$ lies in the plane of the wires and points to the left in Fig. 34-9.

We could have started with wire b, computed the field of induction which it produces at the site of wire a, and then computed the force on wire a. The

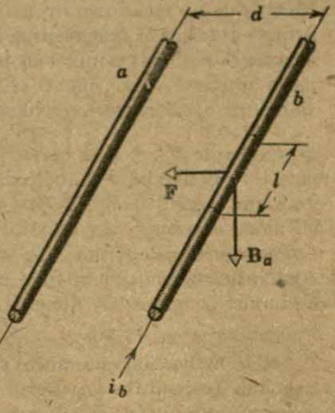


Fig. 34-9 Two parallel wires that carry parallel currents attract each other.

force on wire a would, for parallel currents, point to the right. The forces that the two wires exert on each other are equal and opposite, as they must be according to Newton's law of action and reaction. For antiparallel currents the two wires repel each other.

This discussion reminds us of our discussion of the electric field between two point charges in Section 27-1. There we saw that the charges act on each other through the intermediary of the electric field. The conductors in Fig. 34-9 act on each other through the intermediary of the *magnetic* field. We think in terms of

$$\text{current} \rightleftharpoons \text{field}$$

and not, as in the action-at-a-distance point of view, in terms of

$$\text{current} \rightleftharpoons \text{current}.$$

The attraction between long parallel wires is used to define the ampere. Suppose that the wires are 1 meter apart ($d = 1.0$ meter) and that the two currents are equal ($i_a = i_b = i$). If this common current is adjusted until, by measurement, the force of attraction per unit length between the wires is 2×10^{-7} nt/meter, the current is defined to be 1 *ampere*. From Eq. 34-6,

$$\frac{F}{l} = \frac{\mu_0 i^2}{2\pi d} = \frac{(4\pi \times 10^{-7}\ \text{weber/amp-m})(1\ \text{amp})^2}{(2\pi)(1\ \text{meter})}$$
$$= 2 \times 10^{-7}\ \text{nt/meter}$$

as expected.*

At the National Bureau of Standards primary measurements of current are made with a *current balance*. This consists of a carefully wound coil placed between two other coils, as in Fig. 34-10. The outer pair of coils is fastened to the table, and the inner one is hung from the arm of a balance. The coils are so connected that the current to be measured exists, as a common current, in all three of them.

The coils exert forces on one another—just as the parallel wires of Fig. 34-9 do—which can be measured by loading weights on the balance pan. The current is defined in terms of this measured force and the dimensions of the coils. The current balance is perfectly equivalent to the long parallel wires of Fig. 34-9 but is a much more practical arrangement. Current balance measurements are used primarily to standardize other, more convenient, secondary methods of measuring currents.

▶ **Example 3.** A long horizontal rigidly supported wire carries a current i_a of 100 amp. Directly above it and parallel to it is a fine wire that carries a current i_b of 20 amp and weighs 0.0050 lb/ft (= 0.073 nt/meter). How far above the lower wire should this second wire be strung if we hope to support it by magnetic repulsion?

To provide a repulsion, the two currents must point in opposite directions. For equilibrium, the magnetic force per unit length must equal the weight per unit length and must be oppositely directed.

---

* Note that μ_0 appears in this relation used to define the ampere. As stated on page 848, μ_0 is assigned the (arbitrary) value of $4\pi \times 10^{-7}$ weber/amp-m, and the size of the current that we define as 1 ampere is adjusted to give the required force of attraction per unit length.

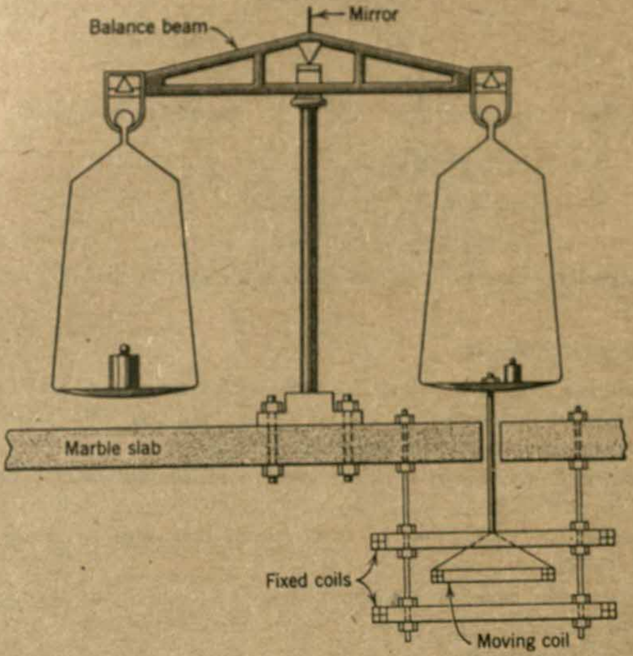


Fig. 34–10 A current balance.

Solving Eq. 34–6 for d yields

$$d = \frac{\mu_0 i_a i_b}{2\pi(F/l)} = \frac{(4\pi \times 10^{-7}\text{ weber/amp-m})(100\text{ amp})(20\text{ amp})}{(2\pi)(0.073\text{ nt/meter})}$$

$$= 5.5 \times 10^{-3}\text{ meter} = 5.5\text{ mm}.$$

We assume that the wire diameters are much smaller than their separation. This assumption is necessary because in deriving Eq. 34–6 we tacitly assumed that the magnetic induction produced by one wire is uniform for all points within the second wire.

Is the equilibrium of the suspended wire stable or unstable against vertical displacements? This can be tested by displacing the wire vertically and examining how the forces on the wire change.

Suppose that the fine wire is suspended *below* the rigidly supported wire. How may it be made to "float"? Is the equilibrium against vertical displacements stable or unstable?

Example 4. Two parallel wires a distance d apart carry equal currents i in opposite directions. Find the magnetic induction for points between the wires and at a distance x from one wire.

Study of Fig. 34–11 shows that $\mathbf{B}_a$ due to the current i_a and $\mathbf{B}_b$ due to the current i_b point in the same direction at P. Each is given by Eq. 34–4 ($B = \mu_0 i/2\pi r$) so that

$$B = B_a + B_b = \frac{\mu_0 i}{2\pi}\left(\frac{1}{x} + \frac{1}{d - x}\right).$$

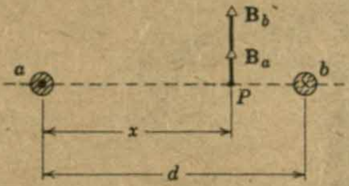


Fig. 34–11 Example 4.

This relationship does not hold for points inside the wires because Eq. 34–4 is not valid there. ◀

34–5 B for a Solenoid

A *solenoid* is a long wire wound in a close-packed helix and carrying a current i. We assume that the helix is very long compared with its diameter. What is the nature of the field of **B** that is set up?

For points close to a single turn of the solenoid, the observer is not aware that the wire is bent in an arc. The wire behaves magnetically almost like a long straight wire, and the lines of **B** due to this single turn are almost concentric circles.

The solenoid field is the vector sum of the fields set up by all the turns that make up the solenoid. Figure 34–12, which shows a "solenoid" with widely spaced turns, suggests that the fields tend to cancel between the wires. It suggests also that, at points inside the solenoid and reasonably far from the wires, **B** is parallel to the solenoid axis. In the limiting case of adjacent square tightly packed wires, the solenoid becomes essentially a cylindrical current sheet and the requirements of symmetry then make the statement just given necessarily true. We assume that it is true in what follows.

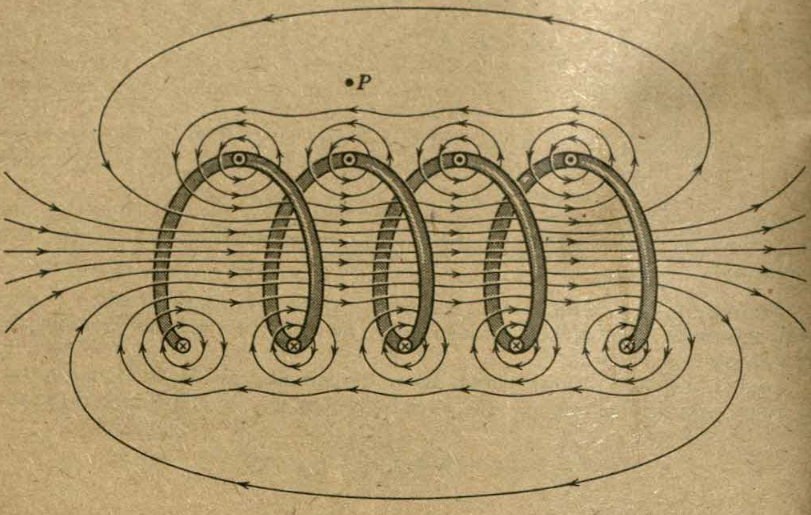


Fig. 34–12 A loosely wound solenoid.

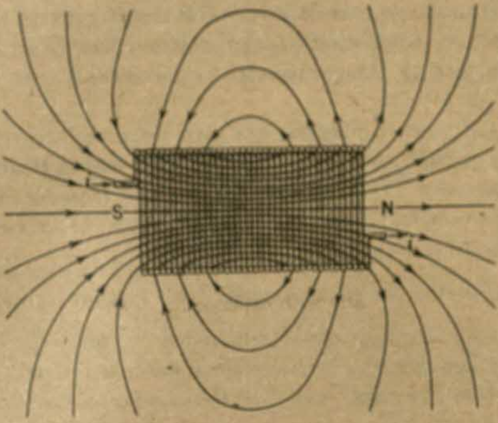


Fig. 34-13 A solenoid of finite length. The right end, from which lines of **B** emerge, behaves like the north pole of a compass needle. The left end behaves like the south pole.

For points such as P in Fig. 34-12 the field set up by the upper part of the solenoid turns (marked ⊙) points to the left and tends to cancel the field set up by the lower part of the solenoid turns (marked ⊗), which points to the right. As the solenoid becomes more and more ideal, that is, as it approaches the configuration of an infinitely long cylindrical current sheet, the field of induction at outside points approaches zero. Taking the external field to be zero is not a bad assumption for a practical solenoid if its length is much greater than its diameter and if we consider only external points near the central region of the solenoid, that is, away from the ends. Figure 34-13 shows the lines of induction for a real solenoid, which is far from ideal in that the length is not much greater than the diameter. Even here the spacing of the lines of induction in the central plane shows that the external field is much weaker than the internal field.

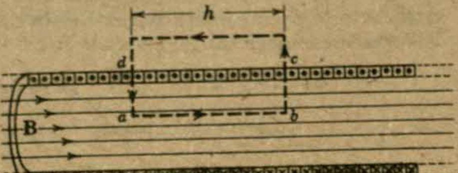
Let us apply Ampère's law,

$$\oint \mathbf{B}\cdot d\mathbf{l} = \mu_0 i,$$

to the rectangular path *abcd* in the ideal solenoid of Fig. 34-14. We write the integral $\oint \mathbf{B}\cdot d\mathbf{l}$ as the sum of four integrals, one for each path segment:

$$\oint \mathbf{B}\cdot d\mathbf{l} = \int_a^b \mathbf{B}\cdot d\mathbf{l} + \int_b^c \mathbf{B}\cdot d\mathbf{l} + \int_c^d \mathbf{B}\cdot d\mathbf{l} + \int_d^a \mathbf{B}\cdot d\mathbf{l}.$$

Fig. 34-14 A section of an ideal solenoid, made of adjacent square turns, equivalent to an infinitely long cylindrical current sheet.



The first integral on the right is Bh, where B is the magnitude of the magnetic induction inside the solenoid and h is the arbitrary length of the path from a to b. Note that path ab, though parallel to the solenoid axis, need not coincide with it.

The second and fourth integrals are zero because for every element of these paths $\mathbf{B}$ is at right angles to the path. This makes $\mathbf{B}\cdot d\mathbf{l}$ zero and thus the integrals are zero. The third integral, which includes the part of the rectangle that lies outside the solenoid, is zero because we have taken $\mathbf{B}$ as zero for all external points for an ideal solenoid.

Thus $\oint \mathbf{B}\cdot d\mathbf{l}$ for the entire rectangular path has the value Bh. The net current i that passes through the area bounded by the path of integration is not the same as the current i_0 in the solenoid because the path of integration encloses more than one turn. Let n be the *number of turns per unit length;* then

$$i = i_0(nh).$$

Ampère's law then becomes

$$Bh = \mu_0 i_0 nh$$

or

$$B = \mu_0 i_0 n. \tag{34-7}$$

Although Eq. 34–7 was derived for an infinitely long ideal solenoid, it holds quite well for actual solenoids for internal points near the center of the solenoid. It shows that B does not depend on the diameter or the length of the solenoid and that B is constant over the solenoid cross section. A solenoid is a practical way to set up a known uniform magnetic field for experimentation, just as a parallel-plate capacitor is a practical way to set up a known uniform electric field.

▶ **Example 5.** A solenoid is 1.0 meter long and 3.0 cm in mean diameter. It has five layers of windings of 850 turns each and carries a current of 5.0 amp.

(*a*) What is B at its center? From Eq. 34–7,

$$B = \mu_0 i_0 n = (4\pi \times 10^{-7}\ \text{weber/amp-m})(5.0\ \text{amp})(5 \times 850\ \text{turns/meter})$$
$$= 2.7 \times 10^{-2}\ \text{weber/meter}^2.$$

We can use Eq. 34–7 even if the solenoid has more than one layer of windings because the diameter of the windings does not enter.

(*b*) What is the magnetic flux Φ_B for a cross section of the solenoid at its center? To the extent that $\mathbf{B}$ is constant, we can calculate the flux from

$$\Phi_B = \int \mathbf{B}\cdot d\mathbf{S} = BA,$$

where A is the effective cross-sectional area. Let us assume that A represents the area of a circular disk whose diameter is the mean diameter of the windings (3.0 cm). The effective area can then be shown to be 7.1×10^{-4} meter², and

$$\Phi_B = BA = (2.7 \times 10^{-2}\ \text{weber/meter}^2)(7.1 \times 10^{-4}\ \text{meter}^2)$$
$$= 1.9 \times 10^{-5}\ \text{weber}.$$

Example 6. *A toroid.* Figure 34-15 shows a toroid, which may be described as a solenoid of finite length bent into the shape of a doughnut. Calculate **B** at interior points.

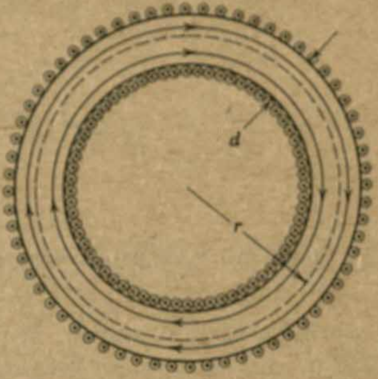


Fig. 34-15 Example 6. A toroid.

From symmetry the lines of **B** form concentric circles inside the toroid, as shown in the figure. Let us apply Ampère's law to the circular path of integration of radius r:

$$\oint \mathbf{B}\cdot d\mathbf{l} = \mu_0 i$$

or

$$(B)(2\pi r) = \mu_0 i_0 N,$$

where i_0 is the current in the toroid windings and N is the total number of turns. This gives

$$B = \frac{\mu_0}{2\pi}\frac{i_0 N}{r}.$$

In contrast to the solenoid, B is not constant over the cross section of a toroid. Show from Ampère's law that B equals zero for points outside an ideal toroid. ◄

34-6 The Biot-Savart Law

Ampère's law can be used to calculate magnetic fields only if the symmetry of the current distribution is high enough to permit the easy evaluation of the line integral $\oint \mathbf{B}\cdot d\mathbf{l}$. This requirement limits the usefulness of the law in practical problems. The law does not fail; it simply becomes difficult to apply in a useful way.

Similarly, in electrostatics, Gauss's law can be used to calculate electric fields only if the symmetry of the charge distribution is high enough to permit the easy evaluation of the surface integral $\oint \mathbf{E}\cdot d\mathbf{S}$. We can, for example, use Gauss's law to find the electric field due to a long uniformly charged rod but we cannot apply it usefully to an electric dipole, for the symmetry is not high enough in this case.

To compute **E** at a given point for an *arbitrary* charge distribution, we divided the distribution into *charge elements dq* and (see Section 27-4) we used Coulomb's law to calculate the field contribution $d\mathbf{E}$ due to each ele-

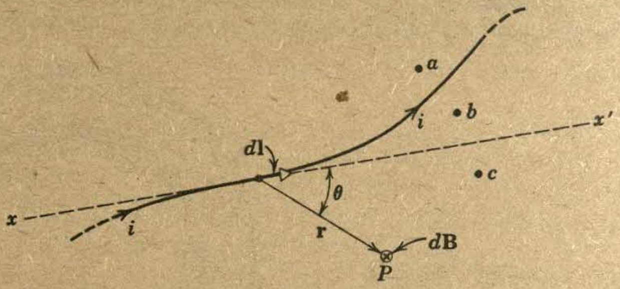


Fig. 34–16 The current element $d\mathbf{l}$ establishes a magnetic field contribution $d\mathbf{B}$ at point P.

ment at the point in question. We found the field **E** at that point by adding, that is, by integrating, the field contributions $d\mathbf{E}$ for the entire distribution.

We now describe a similar procedure for computing **B** at any point due to an arbitrary current distribution. We divide the current distribution into *current elements* and, using the law of Biot and Savart (which we describe below), we calculate the field contribution $d\mathbf{B}$ due to each current element at the point in question. We find the field **B** at that point by integrating the field contributions for the entire distribution.

Figure 34–16 shows an arbitrary current distribution consisting of a current i in a curved wire. The figure also shows a typical current element; it is a length $d\mathbf{l}$ of the conductor carrying a current i. Its direction is that of the tangent to the conductor (dashed line). A current element cannot exist as an isolated entity because a way must be provided to lead the current into the element at one end and out of it at the other. Nevertheless, we can think of an actual circuit as made up of a large number of current elements placed end to end.

Let P be the point at which we want to know the magnetic induction $d\mathbf{B}$ associated with the current element. According to the Biot-Savart law, $d\mathbf{B}$ is given in magnitude by

$$dB = \frac{\mu_0 i}{4\pi}\frac{dl \sin\theta}{r^2}, \tag{34–8}$$

where $\mathbf{r}$ is a displacement vector from the element to P and θ is the angle between this vector and $d\mathbf{l}$. The direction of $d\mathbf{B}$ is that of the vector $d\mathbf{l} \times \mathbf{r}$. In Fig. 34–16, for example, $d\mathbf{B}$ at point P for the current element shown is directed into the page at right angles to the plane of the figure. Note that Eq. 34–8, being an inverse square law that describes the magnetic induction due to a current element, may be viewed as the magnetic equivalent of Coulomb's law, which is an inverse square law that describes the electric field due to a charge element.

The law of Biot and Savart may be written in vector form as

$$d\mathbf{B} = \frac{\mu_0 i}{4\pi}\frac{d\mathbf{l} \times \mathbf{r}}{r^3}. \tag{34–9}$$

This formulation reduces to that of Eq. 34–8 when expressed in terms of magnitudes; it also gives complete information about the direction of $d\mathbf{B}$, namely that it is the same as the direction of the vector $d\mathbf{l} \times \mathbf{r}$.

The resultant field at P is found by integrating Eq. 34–9, or

$$\mathbf{B} = \int d\mathbf{B}, \tag{34–10}$$

where the integral is a vector integral.

▶ **Example 7.** *A long straight wire.* We illustrate the law of Biot and Savart by applying it to find $\mathbf{B}$ due to a current i in a long straight wire. We discussed this problem at length in connection with Ampère's law in Section 34–1.

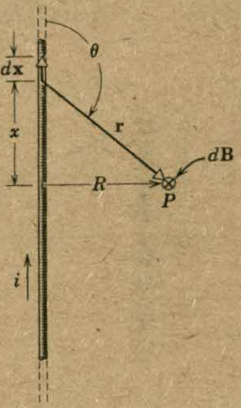


Fig. 34–17 Example 7.

Figure 34–17, a side view of the wire, shows a typical current element $d\mathbf{x}$. The magnitude of the contribution $d\mathbf{B}$ of this element to the magnetic field at P is found from Eq. 34–8, or

$$dB = \frac{\mu_0 i}{4\pi} \frac{dx \sin\theta}{r^2}.$$

The directions of the contributions $d\mathbf{B}$ at point P for all elements are the same, namely, into the plane of the figure at right angles to the page. Thus the vector integral of Eq. 34–10 reduces to a scalar integral, or

$$B = \int dB = \frac{\mu_0 i}{4\pi} \int_{x=-\infty}^{x=+\infty} \frac{\sin\theta\, dx}{r^2}.$$

Now, x, θ, and r are not independent, being related (see Fig. 34–17) by

$$r = \sqrt{x^2 + R^2}$$

and

$$\sin\theta\,[= \sin(\pi - \theta)] = \frac{R}{\sqrt{x^2 + R^2}},$$

so that the expression for B becomes

$$B = \frac{\mu_0 i}{4\pi}\int_{-\infty}^{+\infty}\frac{R\,dx}{(x^2+R^2)^{3/2}}$$

$$= \frac{\mu_0 i}{4\pi R}\left|\frac{x}{(x^2+R^2)^{1/2}}\right|_{x=-\infty}^{x=+\infty}$$

$$= \frac{\mu_0}{2\pi}\frac{i}{R}.$$

This is the result that we arrived at earlier for this problem (see Eq. 34–4). The law of Biot and Savart will always yield results that are consistent with Ampère's law and with experiment.

This problem reminds us of its electrostatic equivalent. We derived an expression for **E** due to a long charged rod, using Gauss's law (Section 28–6); we also solved this problem by integration methods, using Coulomb's law (Section 27–4).

Example 8. *A circular current loop.* Figure 34–18 shows a circular loop of radius R carrying a current i. Calculate **B** for points on the axis.

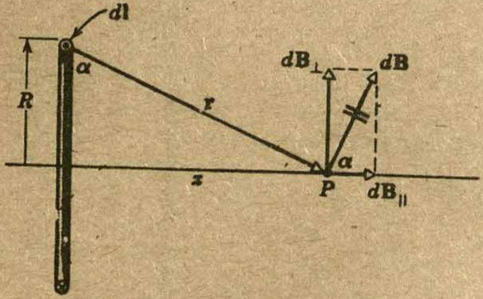


Fig. 34–18 Example 8. A ring of radius R carrying a current i.

The vector $d\mathbf{l}$ for a current element at the top of the loop points perpendicularly out of the page. The angle θ between $d\mathbf{l}$ and $\mathbf{r}$ is 90° and the plane formed by $d\mathbf{l}$ and $\mathbf{r}$ is normal to the page. The vector $d\mathbf{B}$ for this element is at right angles to this plane and thus lies in the plane of the figure and at right angles to $\mathbf{r}$, as the figure shows.

Let us resolve $d\mathbf{B}$ into two components, one, $d\mathbf{B}_\parallel$, along the axis of the loop and another, $d\mathbf{B}_\perp$, at right angles to the axis. Only $d\mathbf{B}_\parallel$ contributes to the total induction **B** at point P. This follows because the components $d\mathbf{B}_\parallel$ for all current elements lie on the axis and add directly; however, the components $d\mathbf{B}_\perp$ point in different directions perpendicular to the axis, and their resultant for the complete loop is zero, from symmetry. Thus

$$B = \int dB_\parallel,$$

where the integral is a simple scalar integration over the current elements.

For the current element shown in Fig. 34–18 we have, from the Biot-Savart law (Eq. 34–8),

$$dB = \frac{\mu_0 i}{4\pi}\frac{dl\sin 90°}{r^2}.$$

We also have
$$dB_{\parallel} = dB \cos \alpha.$$

Combining gives
$$dB_{\parallel} = \frac{\mu_0 i \cos \alpha \, dl}{4\pi r^2}.$$

Figure 34-18 shows that r and α are not independent of each other. Let us express each in terms of a new variable x, the distance from the center of the loop to the point P. The relationships are

$$r = \sqrt{R^2 + x^2}$$

and
$$\cos \alpha = \frac{R}{r} = \frac{R}{\sqrt{R^2 + x^2}}.$$

Substituting these values into the expression for $dB_{\parallel}$ gives

$$dB_{\parallel} = \frac{\mu_0 i R}{4\pi(R^2 + x^2)^{3/2}} \, dl.$$

Note that i, R, and x have the same values for all current elements. Integrating this equation, noting that $\int dl$ is simply the circumference of the loop ($= 2\pi R$), yields

$$B = \int dB_{\parallel} = \frac{\mu_0 i R}{4\pi(R^2 + x^2)^{3/2}} \int dl$$

$$= \frac{\mu_0 i R^2}{2(R^2 + x^2)^{3/2}}. \tag{34-11}$$

◀

If we put $x \gg R$ in Example 8 so that points close to the loop are not considered, Eq. 34-11 reduces to

$$B = \frac{\mu_0 i R^2}{2x^3}.$$

Recalling that πR^2 is the area A of the loop and considering loops with N turns, we can write this equation as

$$B = \frac{\mu_0}{2\pi} \frac{(NiA)}{x^3} = \frac{\mu_0}{2\pi} \frac{\mu}{x^3},$$

where μ is the *magnetic dipole moment* of the current loop. This reminds us of the result derived in Problem 10, Chapter 27 [$E = (1/2\pi\epsilon_0)(p/x^3)$], which is the formula for the *electric* field strength on the axis of an *electric* dipole.

Thus we have shown in two ways that a current loop can be regarded as a magnetic dipole: It experiences a torque given by $\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B}$ when placed in an *external* magnetic field (Eq. 33-11); it generates its own magnetic field given, for points on the axis, by the equation just developed.

Table 34-1 is a summary of the properties of electric and magnetic dipoles.

Table 34–1

SOME DIPOLE EQUATIONS

Property	Dipole Type	Equation
Torque in an external field	electric	$\boldsymbol{\tau} = \mathbf{p}\times\mathbf{E}$
	magnetic	$\boldsymbol{\tau} = \boldsymbol{\mu}\times\mathbf{B}$
Energy in an external field	electric	$U = -\mathbf{p}\cdot\mathbf{E}$
	magnetic	$U = -\boldsymbol{\mu}\cdot\mathbf{B}$
Field at distant points along axis	electric	$E = \frac{1}{2\pi\epsilon_0}\frac{p}{x^3}$
	magnetic	$B = \frac{\mu_0}{2\pi}\frac{\mu}{x^3}$
Field at distant points along perpendicular bisector	electric	$E = \frac{1}{4\pi\epsilon_0}\frac{p}{x^3}$
	magnetic	$B = \frac{\mu_0}{4\pi}\frac{\mu}{x^3}$

▶ **Example 9.** In the Bohr model of the hydrogen atom the electron circulates around the nucleus in a path of radius 5.1×10^{-11} meter at a frequency ν of 6.8×10^{15} rev/sec. (*a*) What value of B is set up at the center of the orbit?

The current is the rate at which charge passes any point on the orbit and is given by

$$i = e\nu = (1.6 \times 10^{-19}\text{ coul})(6.8 \times 10^{15}/\text{sec}) = 1.1 \times 10^{-3}\text{ amp}.$$

B at the center of the orbit is given by Eq. 34–11 with $x = 0$, or

$$B = \frac{\mu_0 i R^2}{2(R^2 + x^2)^{3/2}} = \frac{\mu_0 i}{2R}$$

$$= \frac{(4\pi \times 10^{-7}\text{ weber/amp-m})(1.1 \times 10^{-3}\text{ amp})}{(2)(5.1 \times 10^{-11}\text{ meter})}$$

$$= 14\text{ webers/meter}^2.$$

(*b*) What is the equivalent magnetic dipole moment? From Eq. 33–10,

$$\mu = NiA = (1)(1.1 \times 10^{-3}\text{ amp})(\pi)(5.1 \times 10^{-11}\text{ meter})^2$$

$$= 9.0 \times 10^{-24}\text{ amp-m}^2.$$ ◀

QUESTIONS

1. Can the path of integration around which we apply Ampère's law pass *through* a conductor?

2. Suppose we set up a path of integration around a cable that contains twelve wires with different currents (some in opposite directions) in each wire. How do we calculate i in Ampère's law in such a case?

3. Apply Ampère's law qualitatively to the three paths shown in Fig. 34–19.

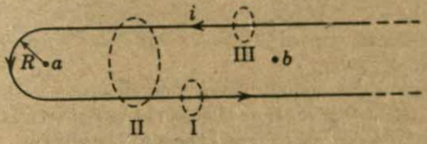


Fig. 34–19

4. Is **B** constant in magnitude for points that lie on a given line of induction?

5. Discuss and compare Gauss's law and Ampère's law.

6. A current is set up in a long copper pipe. Is there a magnetic field (*a*) inside and (*b*) outside the pipe?

7. Equation 34–4 ($B = \mu_0 i/2\pi r$) suggests that a strong magnetic field is set up at points near a long wire carrying a current. Since there is a current i and a magnetic field **B**, why is there not a force on the wire in accord with the equation $\mathbf{F} = i\mathbf{l} \times \mathbf{B}$?

8. In electronics, wires that carry equal but opposite currents are often twisted together to reduce their magnetic effect at distant points. Why is this effective?

9. A beam of 20-Mev protons emerges from a cyclotron. Is a magnetic field associated with these particles?

10. Test the "floating" wire of Example 3 for equilibrium under *horizontal* displacements. Consider that the wire floats above the rigidly supported wire and also below it. Summarize the equilibrium situation for both wire positions and for both vertical and horizontal displacements.

11. Explain qualitatively the forces of interaction between parallel wires carrying parallel or antiparallel currents in terms of Faraday's lines of induction representation.

12. Comment on this statement: "The magnetic induction outside a long solenoid cannot be zero, if only for the reason that the helical nature of the windings produces a field for external points like that of a straight wire along the solenoid axis."

13. A current is sent through a vertical spring from whose lower end a weight is hanging; what will happen?

14. Does Eq. 34–7 ($B = \mu_0 in$) hold for a solenoid of square cross section?

15. What is the direction of the magnetic fields at points a, b, and c in Fig. 34–16 set up by the particular current element shown?

16. In a circular loop of wire carrying a current i, is **B** uniform for all points within the loop?

17. Discuss analogies and differences between Coulomb's law and the Biot-Savart law.

18. Equation 34–9 gives the law of Biot and Savart in vector form. Write its electrostatic equivalent [that is, Eq. 27–6, or $dE = dq/(4\pi\epsilon_0 r^2)$] in vector form.

19. How might you measure the magnetic dipole moment of a compass needle?

20. What is the basis for saying that a current loop is a magnetic dipole?

PROBLEMS

1. A #10 bare copper wire (0.10 in. in diameter) can carry a current of 50 amp without overheating. For this current, what is B at the surface of the wire?

2. A surveyor is using a compass 20 ft below a power line in which there is a steady current of 100 amp. Will this interfere seriously with the compass reading? The horizontal component of the earth's magnetic field at the site is 0.2 gauss.

3. A long straight wire carries a current of 50 amp. An electron, traveling at 10^7 meters/sec, is 5.0 cm from the wire. What force acts on the electron if the electron velocity is directed (*a*) toward the wire, (*b*) parallel to the wire, and (*c*) at right angles to the directions defined by (*a*) and (*b*)?

4. A conductor consists of an infinite number of adjacent wires, each infinitely long and carrying a current i. Show that the lines of **B** will be as represented in Fig. 34–20 and that B for all points in front of the infinite current sheet will be given by

$$B = \tfrac{1}{2}\mu_0 n i,$$

where n is the number of conductors per unit length. Derive both by direct application of Ampère's law and by considering the problem as a limiting case of Example 2.

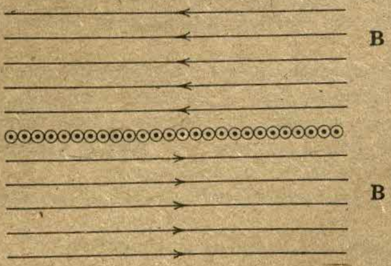


Fig. 34–20

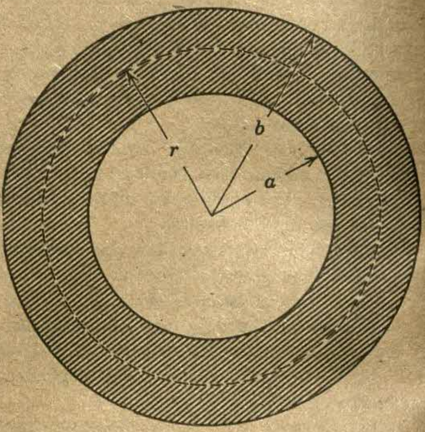


Fig. 34–21

5. Figure 34–21 shows a hollow cylindrical conductor of radii a and b which carries a current i uniformly spread over its cross section. (*a*) Show that the magnetic field B for points inside the body of the conductor (that is, $a < r < b$) is given by

$$B = \frac{\mu_0 i}{2\pi(b^2 - a^2)} \frac{r^2 - a^2}{r}.$$

Check this formula for the limiting case of $a = 0$. (*b*) Make a rough plot of the general behavior of $B(r)$ from $r = 0$ to $r \to \infty$.

6. A long coaxial cable consists of two concentric conductors with the dimensions shown in Fig. 34–22. There are equal and opposite currents i in the conductors. (*a*) Find the magnetic induction B at r within the inner conductor ($r < a$). (*b*) Find B between the two conductors ($a < r < b$). (*c*) Find B within the outer conductor ($b < r < c$). (*d*) Find B outside the cable ($r > c$).

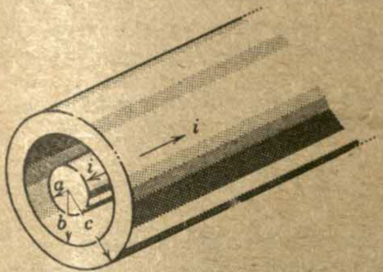


Fig. 34–22

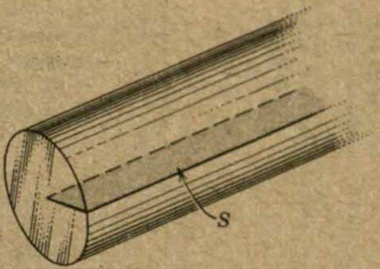


Fig. 34–23

7. A long copper wire carries a current of 10 amp. Calculate the magnetic flux per meter of wire for a plane surface S inside the wire, as in Fig. 34–23.

8. Two long, parallel #10 copper wires (diameter = 0.10 in.) carry currents of 10 amp in opposite directions. (*a*) If their centers are 2.0 cm apart, calculate the flux per meter that exists in the space between the axes of the wires. (*b*) What fraction of the flux in (*a*) lies inside the wires? (*c*) Repeat the calculation of (*a*) for parallel currents.

9. A long wire carrying a current of 100 amp is placed in a uniform external magnetic field of 50 gauss. The wire is at right angles to this external field. Locate the points at which the *resultant* magnetic field is zero.

10. Two long wires a distance d apart carry equal antiparallel currents i, as in Fig. 34–24. Show that B at point P, which is equidistant from the wires, is given by

$$B = \frac{2\mu_0 i d}{\pi(4R^2 + d^2)}.$$

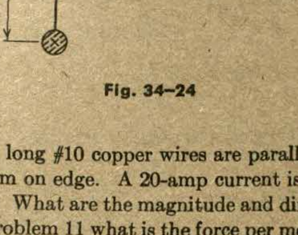


Fig. 34–24

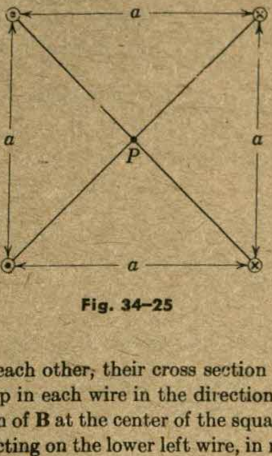


Fig. 34–25

11. Four long #10 copper wires are parallel to each other, their cross section forming a square 20 cm on edge. A 20-amp current is set up in each wire in the direction shown in Fig. 34–25. What are the magnitude and direction of **B** at the center of the square?

12. In Problem 11 what is the force per meter acting on the lower left wire, in magnitude and direction?

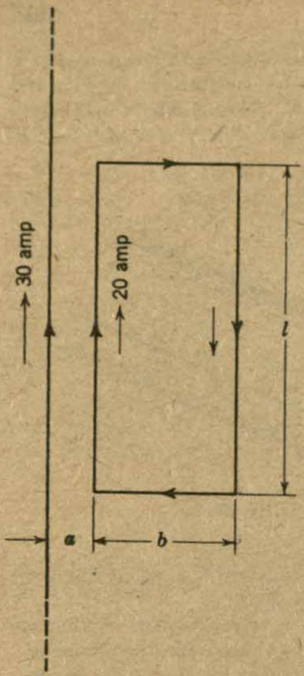


Fig. 34–26

13. Figure 34–26 shows a long wire carrying a current of 30 amp. The rectangular loop carries a current of 20 amp. Calculate the resultant force acting on the loop. Assume that $a = 1.0$ cm, $b = 8.0$ cm, and $l = 30$ cm.

14. Derive the solenoid equation (Eq. 34–7) starting from the expression for the field on the axis of a circular loop (Example 8). (Hint: Subdivide the solenoid into a series of current loops of infinitesimal thickness and integrate.)

15. A square loop of wire of edge a carries a current i. Show that the value of B at the center is given by

$$B = \frac{2\sqrt{2}\,\mu_0 i}{\pi a}.$$

16. A square loop of wire of edge a carries a current i. (a) Show that B for a point on the axis of the loop and a distance x from its center is given by

$$B = \frac{4\mu_0 i a^2}{\pi(4x^2 + a^2)(4x^2 + 2a^2)^{1/2}}.$$

(b) Does this reduce to the result of Problem 15 for $x = 0$? (c) Does the square loop behave like a dipole for points such that $x \gg a$? If so, what is its dipole moment?

17. A straight wire segment of length l carries a current i. (a) Show that the field of induction $\mathbf{B}$ to be associated with this segment, at a distance R from the segment along a perpendicular bisector (see Fig. 34–27), is given in magnitude by

$$B = \frac{\mu_0 i}{2\pi R}\frac{l}{(l^2 + 4R^2)^{1/2}}.$$

b) Does this expression reduce to an expected result as $l \to \infty$?

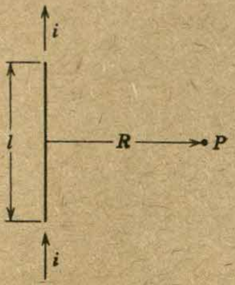


Fig. 34–27

18. The wire shown in Fig. 34–28 carries a current i. What is the magnetic induction at the center C of the semicircle arising from (a) each straight segment of length l, (b) the semicircular segment of radius R, and (c) the entire wire?

19. (a) A wire in the form of a regular polygon of n sides is just enclosed by a circle of radius a. If the current in this wire is i, show that the magnetic induction at the center of the circle is given by

$$B = \frac{\mu_0 n i}{2\pi a}\tan(\pi/n).$$

(b) Show that as $n \to \infty$ this result approaches that of a circular loop.

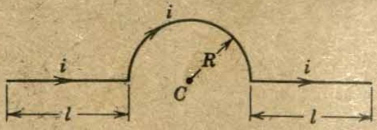


Fig. 34–28

20. (*a*) Show that B at the center of a rectangle of length l and width d, carrying a current i, is given by

$$B = \frac{2\mu_0 i}{\pi}\frac{(l^2 + d^2)^{1/2}}{ld}.$$

(*b*) What does B reduce to for $l \gg d$? Is this a result that you expect?

21. *Helmholtz coils.* Two 300-turn coils are arranged a distance apart equal to their radius, as in Fig. 34–29. For $R = 5.0$ cm and $i = 50$ amp, plot B as a function of distance x along the common axis over the range $x = -5$ cm to $x = +5$ cm, taking $x = 0$ at point P. (Such coils provide an especially uniform field of B near point P.)

22. In Problem 21 let the separation of the coils be a variable z. Show that if z equals R, then not only the first derivative (dB/dx) but also the second (d^2B/dx^2) of B is zero at point P. This accounts for the uniformity of B near point P for this particular coil separation.

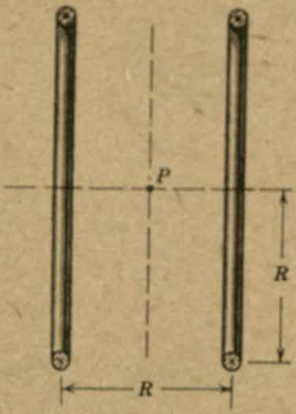


Fig. 34–29

23. A long "hairpin" is formed by bending a piece of wire as shown in Fig. 34–19. If a 10-amp current is set up, what are the direction and magnitude of **B** at point a? At point b? Take $R = 0.50$ cm.

24. Calculate **B** at point P in Fig. 34–30. Assume that $i = 10$ amp and $a = 8.0$ cm.

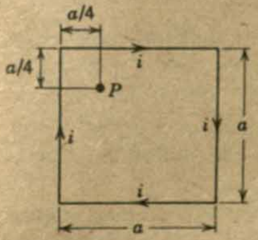


Fig. 34–30

25. A plastic disk of radius R has a charge q uniformly distributed over its surface. If the disk is rotated at an angular frequency ω about its axis, show that (*a*) the induction at the center of the disk is

$$B = \frac{\mu_0\omega q}{2\pi R}$$

and (*b*) the magnetic dipole moment of the disk is

$$\mu = \frac{\omega q R^2}{4}.$$

(Hint: The rotating disk is equivalent to an array of current loops; see Example 8.)

26. You are given a length l of wire in which a current i may be established. The wire may be formed into a circle or a square. Which yields the larger value for B at the central point? See Problem 15.

27. A circular copper loop of radius 10 cm carries a current of 15 amp. At its center is placed a second loop of radius 1.0 cm, having 50 turns and a current of 1.0 amp. (*a*) What magnetic induction **B** does the large loop set up at its center? (*b*) What torque acts on the small loop? Assume that the planes of the two loops are at right angles and that the induction **B** provided by the large loop is essentially uniform throughout the volume occupied by the small loop.

28. Show that it is impossible for a uniform magnetic field **B** to drop abruptly to zero as one moves at right angles to it, as suggested by the horizontal arrow in Fig. 34–31 (see point a). In actual magnets fringing of the lines of force always occurs, which means that **B** approaches zero in a continuous and gradual way. (Hint: Apply Ampère's law to the rectangular path shown by the dashed lines.)

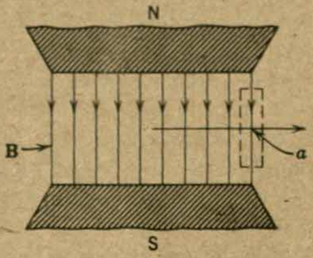


Fig. 34–31

Faraday's Law

CHAPTER 35

35-1 Faraday's Experiments

For some physical laws it is hard to find experiments that lead in a direct and convincing way to the formulation of the law. Gauss's law, for example, emerged only slowly as the common factor with whose aid all electrostatic experiments could be interpreted and correlated. In Chapter 28 we found it best to state Gauss's law first and then to show that the underlying experiments were consistent with it.

Faraday's *law of electromagnetic induction*, which is one of the basic equations of electromagnetism (see Table 38-3), is different in that there are a number of simple experiments from which the law can be—and was—deduced directly. Such experiments were carried out by Michael Faraday in England in 1831 and by Joseph Henry in the United States at about the same time.

Figure 35-1 shows the terminals of a coil connected to a galvanometer. Normally we would not expect this instrument to deflect because there seems to be no electromotive force in this circuit; but if we push a bar magnet toward the coil, with its north pole facing the coil, a remarkable thing happens. *While the magnet is moving*, the galvanometer deflects, showing that a current has been set up in the coil. If the magnet is held stationary with respect to the coil, the galvanometer does not deflect. If the magnet is moved away from the coil, the galvanometer again deflects, but in the opposite direction, which means that the current in the coil is in the opposite direction. If we use the south pole end of a magnet instead of the north pole end, the experiment works as described but the deflections are reversed. Further experimentation shows that *what matters is the relative motion of the*

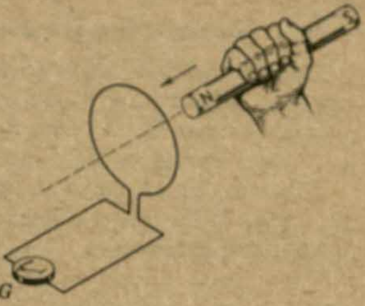


Fig. 35–1 Galvanometer *G* deflects while the magnet is moving with respect to the coil.

magnet and the coil. It makes no difference whether the magnet is moved toward the coil or the coil toward the magnet.

The current that appears in this experiment is called an *induced current* and is said to be set up by an *induced electromotive force.* Faraday was able to deduce from experiments like this the law that gives their magnitude and direction. Such emfs are very important in practice. The chances are good that the lights in the room in which you are reading this book are operated from an induced emf produced in a commercial electric generator.

In another experiment the apparatus of Fig. 35–2 is used. The coils are placed close together but at rest with respect to each other. When the switch S is closed, thus setting up a steady current in the right-hand coil, the galvanometer deflects momentarily; when the switch is opened, thus interrupting this current, the galvanometer again deflects momentarily, but in the opposite direction. No gross objects are moving in this experiment. In Faraday's words:

> When the contact was made, there was a sudden and very slight effect at the galvanometer, and there was also a similar slight effect when the contact with the battery was broken. But whilst the voltaic current was continuing to pass through the one helix, no galvanometrical appearances nor any effect like induction upon the other helix could be perceived, although the active power of the battery was proved to be very great. . . .

Experiment shows that there will be an induced emf in the left coil of Fig. 35–2 whenever the current in the right coil is *changing.* It is the *rate at which the current is changing* and *not the size of the current* that is significant.

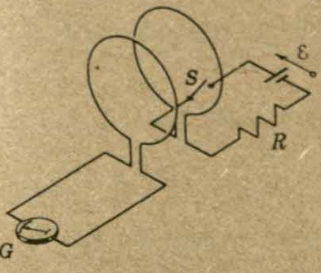


Fig. 35–2 Galvanometer *G* deflects momentarily when switch *S* is closed or opened.

35–2 Faraday's Law of Induction

Faraday had the insight to perceive that the change in the flux Φ_B of magnetic induction for the left coil in the preceding experiments is the important common factor. This flux may be set up by a bar magnet or a current loop. Faraday's law of induction says that the induced emf $\mathcal{E}$ in a circuit is equal to the negative rate at which the flux through the circuit is changing. If the rate of change of flux is in webers/sec, the emf $\mathcal{E}$ will be in volts. In equation form

$$\mathcal{E} = -\frac{d\Phi_B}{dt}. \tag{35–1}$$

This equation is called *Faraday's law of induction*. The minus sign is an indication of the *direction* of the induced emf, a matter we discuss further in Section 35–3.

If Eq. 35–1 is applied to a coil of N turns, an emf appears in every turn and these emfs are to be added. If the coil is so tightly wound that each turn can be said to occupy the same region of space, the flux through each turn will then be the same. The flux through each turn is also the same for (ideal) toroids and solenoids (see Section 34–5). The induced emf in all such devices is given by

$$\mathcal{E} = -N\frac{d\Phi_B}{dt} = -\frac{d(N\Phi_B)}{dt}, \tag{35–2}$$

where $N\Phi_B$ measures the so-called *flux linkages* in the device.

Figures 35–1 and 35–2 suggest that there are at least two ways in which we can make the flux through a circuit change and thus induce an emf in that circuit. The coil that is connected to the galvanometer cannot tell in which of these experiments it is participating; it is aware only that the flux passing through its cross-sectional area is changing. The flux through a circuit can also be changed by changing its shape, that is, by squeezing or stretching it.

▶ **Example 1.** A long solenoid has 200 turns/cm and carries a current of 1.5 amp; its diameter is 3.0 cm. At its center we place a 100-turn, close-packed coil of diameter 2.0 cm. This coil is arranged so that **B** at the center of the solenoid is parallel to its axis. The current in the solenoid is reduced to zero and then raised to 1.5 amp in the other direction at a steady rate over a period of 0.050 sec. What induced emf appears in the coil while the current is being changed?

The induction B at the center of the solenoid is given by Eq. 34–7, or

$$\begin{aligned} B = \mu_0 ni &= (4\pi \times 10^{-7}\ \text{weber/amp-m})(200 \times 10^2\ \text{turns/meter})(1.5\ \text{amp}) \\ &= 3.8 \times 10^{-2}\ \text{weber/meter}^2. \end{aligned}$$

The area of the coil (not of the solenoid) is 3.1×10^{-4} meter^2. The initial flux Φ_B through each turn of the coil is given by

$$\Phi_B = BA = (3.8 \times 10^{-2}\ \text{weber/meter}^2)(3.1 \times 10^{-4}\ \text{meter}^2) = 1.2 \times 10^{-5}\ \text{weber}.$$

The flux goes from an initial value of 1.2×10^{-5} weber to a final value of -1.2×10^{-5} weber. The *change* in flux $\Delta\Phi_B$ for each turn of the coil during the 0.050-sec period is thus twice the initial value. The induced emf is given by

$$\mathcal{E} = -\frac{N\Delta\Phi_B}{\Delta t} = -\frac{(100)(2 \times 1.2 \times 10^{-5}\text{ weber})}{0.050\text{ sec}} = -4.8 \times 10^{-2}\text{ volt} = -48\text{ mv}.$$

The minus sign deals with the *direction* of the emf, as we explain below. ◀

35-3 Lenz's Law

So far we have not specified the directions of the induced emfs. Although these directions can be found from a formal analysis of Faraday's law, we prefer to find them from the conservation-of-energy principle which, in this context, takes the form of Lenz's law: *The induced current will appear in such a direction that it opposes the change that produced it.** The minus sign in Faraday's law suggests this opposition. In mechanics the energy principle often allows us to draw conclusions about mechanical systems without analyzing them in detail. We use the same approach here.

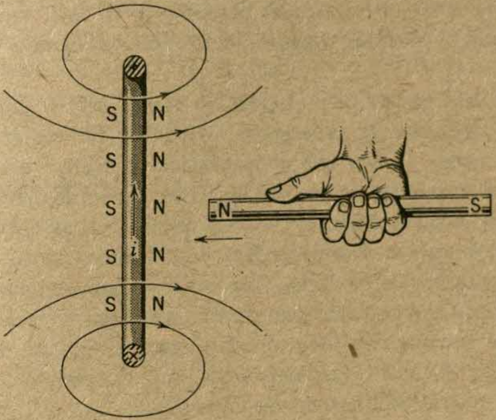


Fig. 35-3 If the magnet is moved toward the loop, the induced current points as shown, setting up a magnetic field that opposes the motion of the magnet.

Lenz's law refers to induced *currents*, which means that it applies only to closed circuits. If the circuit is open, we can usually think in terms of what would happen if it were closed and in this way find the direction of the induced emf.

Consider the first of Faraday's experiments described in Section 35-1. Figure 35-3 shows the north pole of a magnet and a cross section of a nearby conducting loop. As we push the magnet toward the loop, an induced current is set up in the loop. What is its direction?

---

* Heinrich Friedrich Lenz (1804-1865) deduced this law in 1834. Faraday also discovered how to determine the directions of induced emfs, but he did not express his results as succinctly as Lenz.

A current loop sets up a magnetic field at distant points like that of a magnetic dipole, one face of the loop being a north pole, the opposite face being a south pole. The north pole, as for bar magnets, is that face *from* which the lines of **B** emerge. If, as Lenz's law predicts, the loop in Fig. 35–3 is to oppose the motion of the magnet toward it, the face of the loop toward the magnet must become a north pole. The two north poles—one of the current loop and one of the magnet—will repel each other. The right-hand rule shows that for the magnetic field set up by the loop to emerge from the right face of the loop the induced current must be as shown. The current will be counterclockwise as we sight along the magnet toward the loop.

When we push the magnet toward the loop (or the loop toward the magnet), an induced current appears. In terms of Lenz's law this pushing is the "change" that produces the induced current, and, according to this law, the induced current will oppose the "push." If we pull the magnet away from the coil, the induced current will oppose the "pull" by creating a *south* pole on the right-hand face of the loop of Fig. 35–3. To make the right-hand face a south pole, the current must be opposite to that shown in Fig. 35–3. Whether we pull or push the magnet, its motion will always be automatically opposed.

The agent that causes the magnet to move, either toward the coil or away from it, will always experience a resisting force and will thus be required to do work. From the conservation-of-energy principle this work done on the system must be exactly equal to the Joule heat produced in the coil, since these are the only two energy transfers that occur in the system. If we move the magnet more rapidly, we will have to do work at a faster rate and the rate of the Joule heating will increase correspondingly. If we cut the loop and then perform the experiment, there will be no induced current, no Joule heating, no force on the magnet, and no work required to move it. There will still be an emf in the loop, but, like a battery connected to an open circuit, it will not set up a current.

If the current in Fig. 35–3 were in the *opposite* direction to that shown, the face of the loop toward the magnet would be a south pole, which would pull the bar magnet toward the loop. We would only need to push the magnet slightly to start the process and then the action would be self-perpetuating. The magnet would accelerate toward the loop, increasing its kinetic energy all the time. At the same time Joule heat would appear in the loop at a rate that would increase with time. This would indeed be a something-for-nothing situation! Needless to say, it does not occur.

Let us apply Lenz's law to Fig. 35–3 in a different way. Figure 35–4 shows the lines of **B** for the bar magnet.* On this point of view the "change" is the increase in Φ_B through the loop caused by bringing the magnet nearer. The induced current opposes this change by setting up a field that tends to oppose the increase in flux caused by the moving magnet. Thus the field

---

* There are two fields of **B** in this problem—one connected with the current loop and one with the bar magnet. The student must always be certain which one is meant.

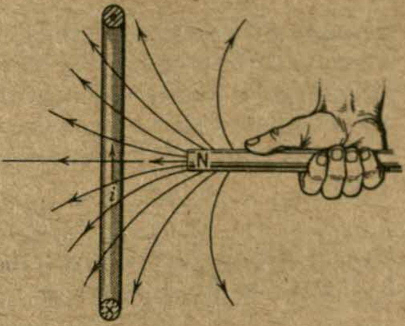


Fig. 35-4 In moving the magnet toward the loop, we increase Φ_B through the loop.

due to the induced current must point from left to right through the plane of the coil, in agreement with our earlier conclusion.

It is not significant here that the induced field opposes the magnet *field* but rather that it opposes the *change*, which in this case is the *increase* in Φ_B through the loop. If we withdraw the magnet, we reduce Φ_B through the loop. The induced field will now oppose this decrease in Φ_B (that is, the change) by *re-enforcing* the magnet field. In each case the induced field opposes the change that gives rise to it.

35-4 Induction—A Quantitative Study

The example of Fig. 35-4, although easy to understand qualitatively, does not lend itself to quantitative calculations. Consider then Fig. 35-5, which shows a rectangular loop of wire of width l, one end of which is in a uniform field **B** pointing at right angles to the plane of the loop. This field of **B** may be produced in the gap of a large electromagnet like that of Fig. 33-2. The dashed lines show the assumed limits of the magnetic field. The experiment consists in pulling the loop to the right at a constant speed v.

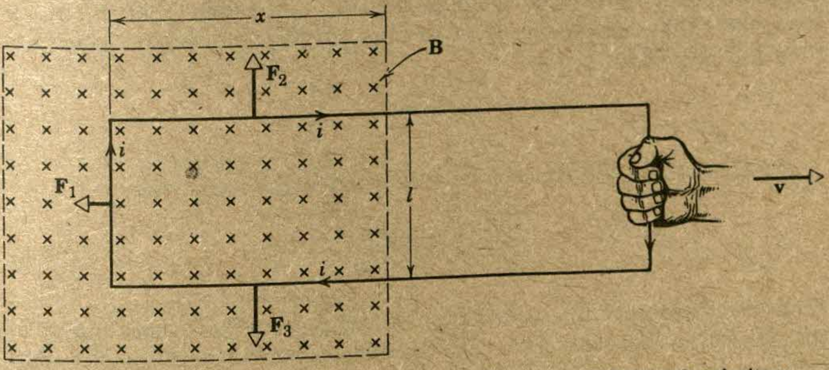


Fig. 35-5 A rectangular loop is pulled out of a magnetic field with velocity v.

Note that the situation described by Fig. 35–5 does not differ in any essential particular from that of Fig. 35–4. In each case a conducting loop and a magnet are in relative motion; in each case the flux of the field of the magnet through the loop is being caused to change with time.

The flux Φ_B enclosed by the loop in Fig. 35–5 is

$$\Phi_B = Blx,$$

where lx is the area of that part of the loop in which B is not zero. The emf $\mathcal{E}$ is found from Faraday's law, or

$$\mathcal{E} = -\frac{d\Phi_B}{dt} = -\frac{d}{dt}(Blx) = -Bl\frac{dx}{dt} = Blv, \tag{35-3}$$

where we have set $-dx/dt$ equal to the speed v at which the loop is pulled out of the magnetic field. Note that the only dimension of the loop that enters into Eq. 35–3 is the length l of the left end conductor. As we shall see later, the induced emf in Fig. 35–5 may be regarded as localized here. An induced emf such as this, produced by pulling a conductor through a magnetic field, is sometimes called a *motional emf.*

The emf Blv sets up a current in the loop given by

$$i = \frac{\mathcal{E}}{R} = \frac{Blv}{R}, \tag{35-4}$$

where R is the loop resistance. From Lenz's law, this current (and thus $\mathcal{E}$) must be clockwise in Fig. 35–5; it opposes the "change" (the decrease in Φ_B) by setting up a field that is parallel to the external field within the loop.

The current in the loop will cause forces $\mathbf{F}_1$, $\mathbf{F}_2$, and $\mathbf{F}_3$ to act on the three conductors, in accord with Eq. 33–6a, or

$$\mathbf{F} = i\mathbf{l} \times \mathbf{B}. \tag{35-5}$$

Because $\mathbf{F}_2$ and $\mathbf{F}_3$ are equal and opposite, they cancel each other; $\mathbf{F}_1$, which is the force that opposes our effort to move the loop, is given in magnitude from Eqs. 35–5 and 35–4 as

$$F_1 = ilB \sin 90° = \frac{B^2l^2v}{R}.$$

The agent that pulls the loop must do work at the steady rate of

$$P = F_1v = \frac{B^2l^2v^2}{R}. \tag{35-6}$$

From the principle of the conservation of energy, Joule heat must appear in the resistor at this same rate. We introduced the conservation-of-energy principle into our derivation when we wrote down the expression for the current (Eq. 35–4); the student will recall that the relation $i = \mathcal{E}/R$ for single-loop circuits is a direct consequence of this principle. Thus we should be able to write down the expression for the rate of Joule heating in the loop with the expectation that we will obtain a result identical with Eq. 35–6. Recalling Eq. 35–4, we put

$$P_J = i^2R = \left(\frac{Blv}{R}\right)^2 R = \frac{B^2l^2v^2}{R},$$

which is indeed the expected result. This example provides a quantitative illustration of the conversion of mechanical energy (the work done by an external agent) into electrical energy (the induced emf) and finally into thermal energy (the Joule heating).

Figure 35–6 shows a side view of the coil in the field. In Fig. 35–6*a* the coil is stationary; in Fig. 35–6*b* we are moving it to the right; in Fig. 35–6*c* we are moving it to the left. The lines of induction in these figures represent the *resultant field* produced by the vector addition of the field $\mathbf{B}_0$ due to the magnet and the field $\mathbf{B}_i$ due to the induced current, if any, in the coil.

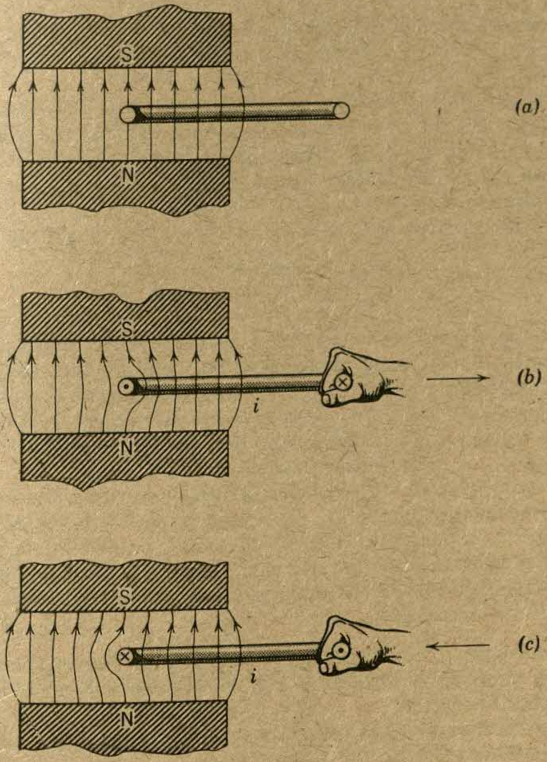


Fig. 35–6 Side view of a rectangular loop in a magnetic field showing the loop (*a*) at rest, being pulled out, and (*c*) being pushed in.

These lines suggest convincingly that the agent moving the coil always experiences an opposing force.

▶ **Example 2.** Figure 35–7 shows a rectangular loop of resistance R, width l, and length a being pulled at constant speed v through a region of thickness d in which a uniform field of induction **B** is set up by a magnet.

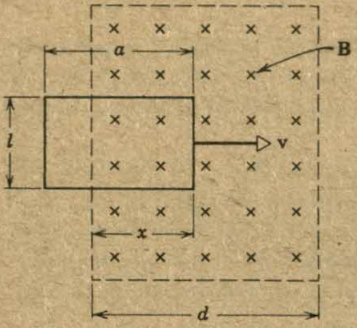


Fig. 35–7 Example 2. A rectangular loop is caused to move with a velocity **v** through a magnetic field. The position of the loop is measured by x, the distance between the effective left edge of the field **B** and the right end of the loop.

(*a*) Plot the flux Φ_B through the loop as a function of the coil position x. Assume that $l = 4$ cm, $a = 10$ cm, $d = 15$ cm, $R = 16$ ohms, $B = 2.0$ webers/meter^2, and $v = 1.0$ meter/sec.

The flux Φ_B is zero when the loop is not in the field; it is Bla when the loop is entirely in the field; it is Blx when the loop is entering the field and $Bl[a - (x - d)]$ when the loop is leaving the field. These conclusions, which the student should verify, are shown graphically in Fig. 35–8*a*.

(*b*) Plot the induced emf $\mathcal{E}$.

The induced emf $\mathcal{E}$ is given by $\mathcal{E} = -d\Phi_B/dt$, which can be written as

$$\mathcal{E} = -\frac{d\Phi_B}{dt} = -\frac{d\Phi_B}{dx}\frac{dx}{dt} = -\frac{d\Phi_B}{dx}v,$$

where $d\Phi_B/dx$ is the slope of the curve of Fig. 35–8*a*. $\mathcal{E}(x)$ is plotted in Fig. 35–8*b*. Lenz's law, from the same type of reasoning as that used for Fig. 35–5, shows that when the coil is entering the field the emf $\mathcal{E}$ acts counterclockwise as seen from above. Note that there is no emf when the coil is entirely in the magnetic field because the flux Φ_B through the coil is not changing with time, as Fig. 35–8*a* shows.

(*c*) Plot the rate P of Joule heating in the loop.

This is given by $P = \mathcal{E}^2/R$. It may be calculated by squaring the ordinate of the curve of Fig. 35–8*b* and dividing by R. The result is plotted in Fig. 35–8*c*.

If the fringing of the magnetic field, which cannot be avoided in practice (see Problem 34–28), is taken into account, the sharp bends and corners in Fig. 35–8 will be replaced by smooth curves. What changes would occur in the curves of Fig. 35–8 if the coil were open circuited?

Example 3. A copper rod of length L rotates at angular frequency ω in a uniform field of magnetic induction **B** as shown in Fig. 35–9. Find the emf $\mathcal{E}$ developed between the two ends of the rod.

If a wire of length dl is moved at velocity **v** at right angles to a field **B**, a motional emf $d\mathcal{E}$ will be developed (see Eq. 35–3) given by

$$d\mathcal{E} = Bv\,dl.$$

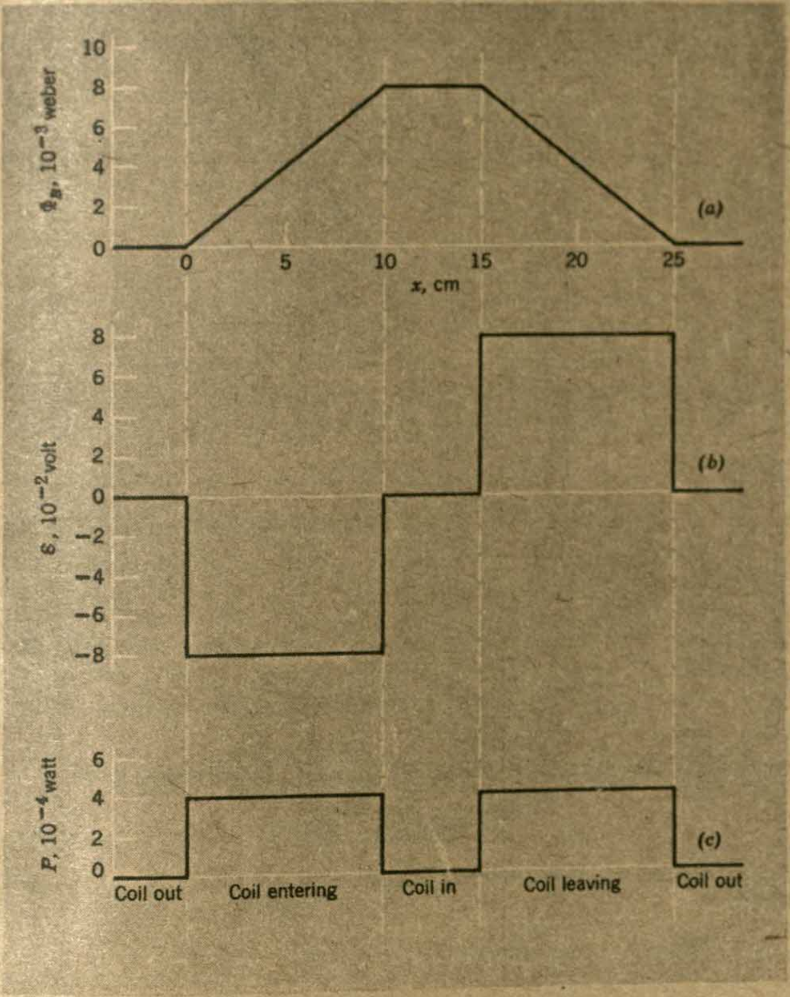


Fig. 35-8 Example 2.

The rod of Fig. 35-9 may be divided into elements of length dl, the linear speed v of each element being ωl. Each element is perpendicular to **B** and is also moving in a direction at right angles to **B** so that

$$\mathcal{E} = \int d\mathcal{E} = \int_0^L Bv\,dl = \int_0^L B(\omega l)\,dl = \tfrac{1}{2}B\omega L^2.$$

For a second approach, consider that at any instant the flux enclosed by the sector aOb in Fig. 35-9 is given by

$$\Phi_B = BA = B(\tfrac{1}{2}L^2\theta),$$

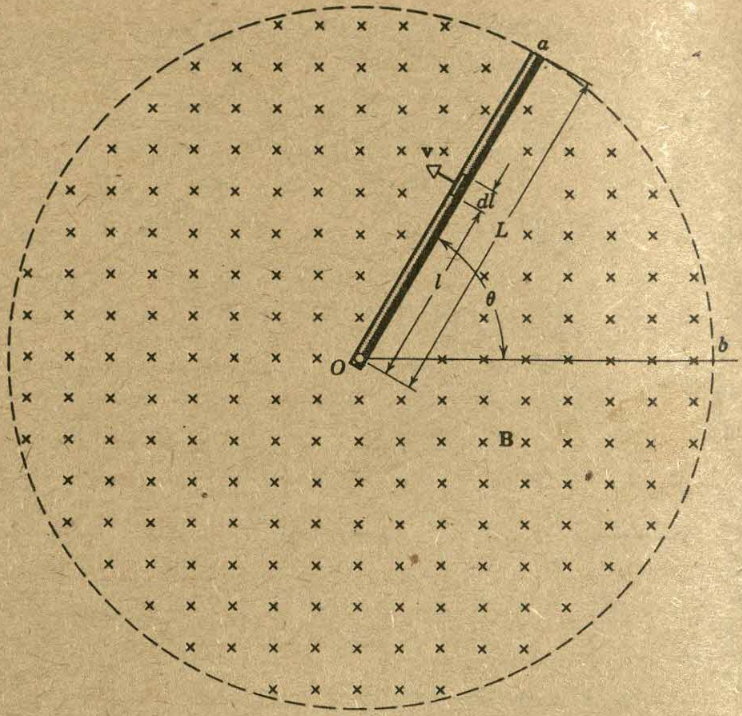


Fig. 35–9 Example 3.

where $\frac{1}{2}L^2\theta$ can be shown to be the area of the sector. Differentiating gives

$$\frac{d\Phi_B}{dt} = \frac{1}{2}BL^2\frac{d\theta}{dt} = \frac{1}{2}B\omega L^2.$$

From Faraday's law, this is precisely the magnitude of $\mathcal{E}$, which agrees with the result just derived. ◀

35–5 Time-Varying Magnetic Fields

So far we have considered emfs induced by the relative motion of magnets and coils. In this section we assume that there is no physical motion of gross objects but that the magnetic field may vary with time. If a conducting loop is placed in such a time-varying field, the flux through the loop will change and an induced emf will appear in the loop. This emf will set the charge carriers in motion, that is, it will induce a current.

From a microscopic point of view we can say, equally well, that the changing flux of **B** sets up an induced electric field **E** at various points around the loop. These induced electric fields are just as real as electric fields set up by static charges and will exert a force **F** on a test charge q_0 given by $\mathbf{F} = q_0\mathbf{E}$. Thus we can restate Faraday's law of induction in a loose but informative way as: *A changing magnetic field produces an electric field.*

To fix these ideas, consider Fig. 35–10, which shows a uniform field of induction **B** at right angles to the plane of the page. We assume that **B** is increasing in magnitude at the same constant rate dB/dt at every point. This could be done by causing the current in the windings of the electromagnet that establishes the field to increase with time in the proper way.

The circle of arbitrary radius r shown in Fig. 35–10 encloses, at any instant, a flux Φ_B. Because this flux is changing with time, an induced emf given by $\mathcal{E} = -d\Phi_B/dt$ will appear around the loop. The electric fields **E** induced at various points of the loop must, from symmetry, be tangent to the loop. Thus the electric lines of force that are set up by the changing magnetic field are in this case concentric circles.

If we consider a test charge q_0 moving around the circle of Fig. 35–10, the work W done on it per revolution is, in terms of the definition of an emf, simply $\mathcal{E}q_0$. From another point of view, it is $(q_0E)(2\pi r)$, where q_0E is the force that acts on the charge and $2\pi r$ is the distance over which the force acts. Setting the two expressions for W equal and canceling q_0 yields

$$\mathcal{E} = E2\pi r. \tag{35–7}$$

In a more general case than that of Fig. 35–10 we must write

$$\mathcal{E} = \oint \mathbf{E}\cdot d\mathbf{l}. \tag{35–8}$$

If this integral is evaluated for the conditions of Fig. 35–10, we obtain Eq. 35–7 at once. If Eq. 35–8 is combined with Eq. 35–1 ($\mathcal{E} = -d\Phi_B/dt$), Faraday's law of induction can be written as

$$\oint \mathbf{E}\cdot d\mathbf{l} = -\frac{d\Phi_B}{dt}, \tag{35–9}$$

which is the form in which this law is expressed in Table 38–3.

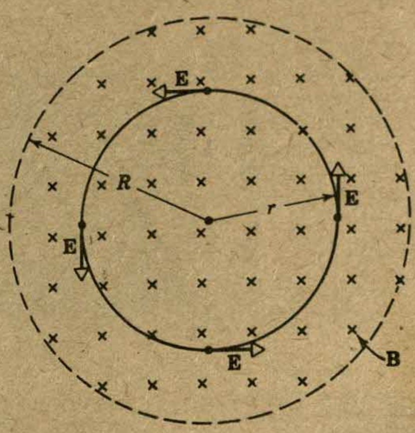


Fig. 35–10 The induced electric fields at four points produced by an increasing magnetic field. We assume that B has symmetry about an axis perpendicular to the page through the center of the circle of radius r. This can be arranged by assuming that the magnetic induction decreases slightly along a radius starting from the center of the figure, the value of B being constant for a particular value of r. (The magnetic field cannot end abruptly at radius R—unless a special arrangement of currents is provided—but must approach zero gradually. This "fringing" does not change any of the arguments of this section.

▶ **Example 4.** Let B in Fig. 35–10 be increasing at the rate dB/dt. Let R be the radius of the cylindrical region in which the magnetic field is assumed to exist. What is the magnitude of the electric field **E** at any radius r? Assume that $dB/dt =$ 0.10 weber/m²-sec and $R = 10$ cm.

(*a*) For $r < R$, the flux Φ_B through the loop is

$$\Phi_B = B(\pi r^2).$$

Substituting into Faraday's law (Eq. 35–9),

$$\oint \mathbf{E}\cdot d\mathbf{l} = -\frac{d\Phi_B}{dt}$$

yields $$(E)(2\pi r) = -\frac{d\Phi_B}{dt} = -(\pi r^2)\frac{dB}{dt}.$$

Solving for E yields $$E = -\tfrac{1}{2}r\frac{dB}{dt}.$$

The minus sign is retained to suggest that the induced electric field **E** acts to *oppose* the change of the magnetic field. Note that $E(r)$ depends on dB/dt and not on B. Substituting numerical values, assuming $r = 5$ cm, yields, for the magnitude of E,

$$E = \tfrac{1}{2}r\frac{dB}{dt} = (\tfrac{1}{2})(5\times 10^{-2}\text{ meter})\left(\frac{0.10\text{ weber}}{\text{m}^2\text{-sec}}\right) = 2.5\times 10^{-3}\text{ volt/meter}.$$

(*b*) For $r > R$ the flux through the loop is

$$\Phi_B = \int \mathbf{B}\cdot d\mathbf{S} = B(\pi R^2).$$

This equation is true because $\mathbf{B}\cdot d\mathbf{S}$ is zero for those points of the loop that lie outside the effective boundary of the magnetic field.

From Faraday's law (Eq. 35–9),

$$(E)(2\pi r) = -\frac{d\Phi_B}{dt} = -(\pi R^2)\frac{dB}{dt}.$$

Solving for E yields $$E = -\frac{1}{2}\frac{R^2}{r}\frac{dB}{dt}.$$

These two expressions for $E(r)$ yield the same result, as they must, for $r = R$. Figure 35–11 is a plot of the magnitude of $E(r)$ for the numerical values given. ◀

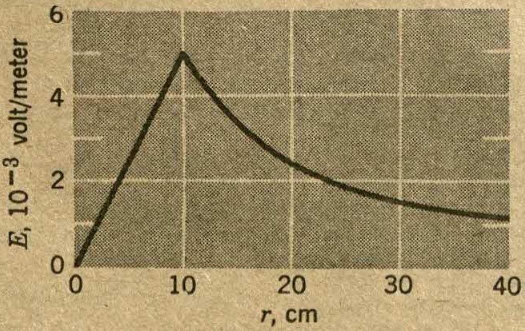


Fig. 35–11 Example 4. If the fringing of the field in Fig. 35–10 were to be taken into account, the result would be a rounding of the sharp cusp at $r = R$ (= 10 cm).

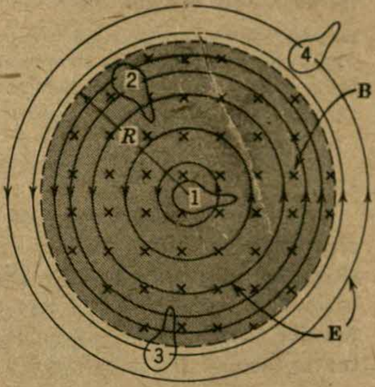


Fig. 35-12 Showing the circular lines of **E** from an increasing magnetic field. The four loops are imaginary paths around which an emf can be calculated.

In applying Lenz's law to Fig. 35-10, imagine that a circular conducting loop is placed concentrically in the field. Since Φ_B through this loop is increasing, the induced current in the loop will tend to oppose this "change" by setting up a magnetic field of its own that points up within the loop. Thus the induced current i must be counterclockwise, which means that the lines of the induced electric field **E**, which is responsible for the current, must also be counterclockwise. If the magnetic field in Fig. 35-10 were *decreasing* with time, the induced current and the lines of force of the induced electric field **E** would be clockwise, again opposing the *change* in Φ_B.

Figure 35-12 shows four of many possible loops to which Faraday's law may be applied. For loops 1 and 2, the induced emf $\mathcal{E}$ is the same because these loops lie entirely within the changing magnetic field and thus have the same value of $d\Phi_B/dt$. Note that even though the emf $\mathcal{E}$ $(= \oint \mathbf{E}\cdot d\mathbf{l})$ is the same for these two loops, the distribution of electric fields **E** around the perimeter of each loop, as indicated by the electric lines of force, is different. For loop 3 the emf is less because Φ_B and $d\Phi_B/dt$ for this loop are less, and for loop 4 the induced emf is zero.

The induced electric fields that are set up by the induction process are not associated with charges but with a changing magnetic flux. Although both kinds of electric fields exert forces on charges, there is a difference between them. The simplest manifestation of this difference is that lines of **E** associated with a changing magnetic flux can form closed loops (see Fig. 35-12); lines of **E** associated with charges cannot but can always be drawn to start on a positive charge and end on a negative charge.

Equation 29-5, which defined the potential difference between two points a and b, is

$$V_b - V_a = \frac{W_{ab}}{q_0} = -\int_a^b \mathbf{E}\cdot d\mathbf{l}.$$

We have insisted that if potential is to have any useful meaning this integral (and W_{ab}) must have the same value for every path connecting a with b. This proved to be true for every case examined in earlier chapters.

An interesting special case comes up if a and b are the same point. The path connecting them is now a closed loop; V_a must be identical with V_b and this equation reduces to

$$\oint \mathbf{E}\cdot d\mathbf{l} = 0. \tag{35-10}$$

However, when changing magnetic flux is present, $\oint \mathbf{E} \cdot d\mathbf{l}$ is precisely *not* zero but is, according to Faraday's law (see Eq. 35–9), $-d\Phi_B/dt$. Electric fields associated with stationary charges are *conservative*, but those associated with changing magnetic fields are *nonconservative*; see Section 8–2. Electric potential, which can be defined only for a conservative force, *has no meaning for electric fields produced by induction.*

35–6 The Betatron

The betatron is a device used to accelerate electrons to high speeds by allowing them to be acted upon by induced electric fields that are set up by a changing magnetic flux. It provides an excellent illustration of the "reality" of such induced fields. The energetic electrons can be used for fundamental research in physics or to produce penetrating X-rays which are useful in cancer therapy and in industry.

Figure 35–13 shows the 100-Mev betatron at the General Electric Company. At this energy the electron speed is 0.999986*c*, where *c* is the speed of light, so that relativistic mechanics must certainly be used in the analysis of its operation. Figure 35–14 shows a vertical cross section through the central part of the betatron to which the man in Fig. 35–13 is pointing.

The magnetic field in the betatron has several functions: (*a*) it guides the electrons in a circular path; (*b*) it accelerates the electrons in this path;

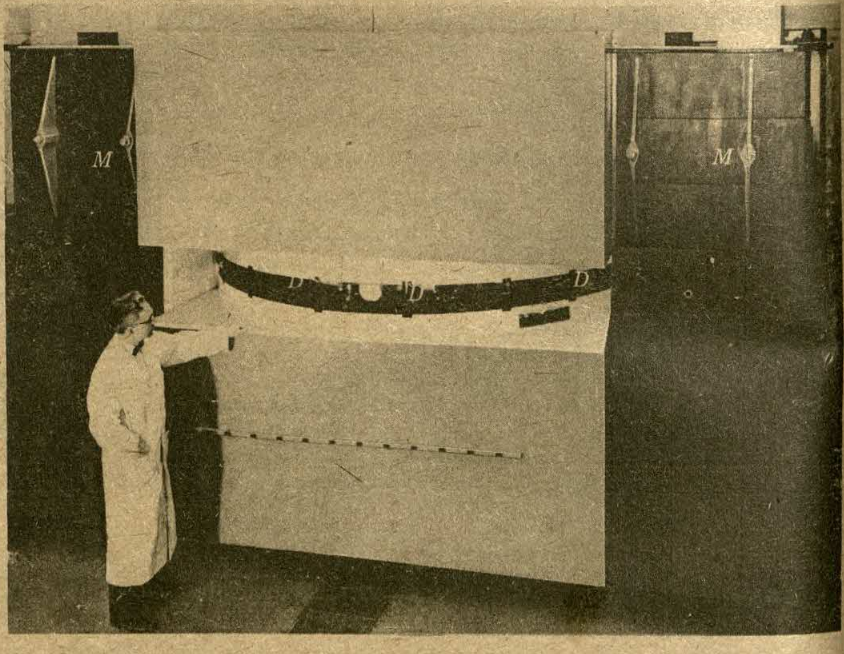


Fig. 35–13 A 100-Mev betatron. *M* shows the magnet, *C* the magnetizing coils, and *D* the region in which the "doughnut" is located. (Courtesy General Electric Company.)

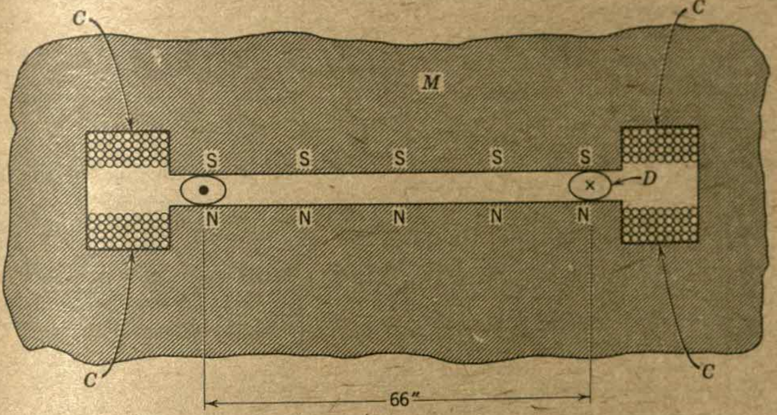


Fig. 35–14 Cross section of a betatron, showing magnet *M*, coils *C*, and "doughnut" *D*. Electrons emerge from the page at the left and enter it at the right.

(*c*) it keeps the radius of the orbit in which the electrons are moving a constant; (*d*) it introduces the electrons into the orbit initially and removes them from the orbit after they have reached full energy; and finally (*e*) it provides a restoring force that resists any tendency for the electrons to leave their orbit, either vertically or radially. It is remarkable that it is possible to do all these things by proper shaping and control of the magnetic field.

The object marked *D* in Fig. 35–14 is an evacuated glass "doughnut" inside which the electrons travel. Their orbit is a circle at right angles to the plane of the figure. The electrons emerge from the plane at the left (·) and enter it at the right (×). In the General Electric machine the radius of the electron path is 33 in. The coils *C* and the 130-ton steel magnet shown in Fig. 35–13 provide the magnetic flux that passes through the plane of this orbit.

The current in coils *C* is made to alter periodically, 60 times/sec, to produce a changing flux through the orbit, shown in Fig. 35–15. Here Φ_B is taken as positive when **B** is pointing up, as in Fig. 35–14. If the electrons are to circulate in the direction shown, they must do so during the positive half-cycle, marked *ac* in Fig. 35–15. The student should verify this (see Section 33–6). The electrons are accelerated by electric fields set up by the changing flux. The direction of these induced fields depends on the sign of $d\Phi_B/dt$ and must be chosen to accelerate, and not to decelerate, the electrons. Thus only half the positive half-cycle in Fig. 35–15 can be used for acceleration; it will prove to be *ab*.

The average value of $d\Phi_B/dt$ during the quarter-cycle *ab* is the slope of the dashed line, or

$$\overline{\frac{d\Phi_B}{dt}} = \frac{1.8 \text{ weber}}{4.2 \times 10^{-3} \text{ sec}} = 430 \text{ volts.}$$

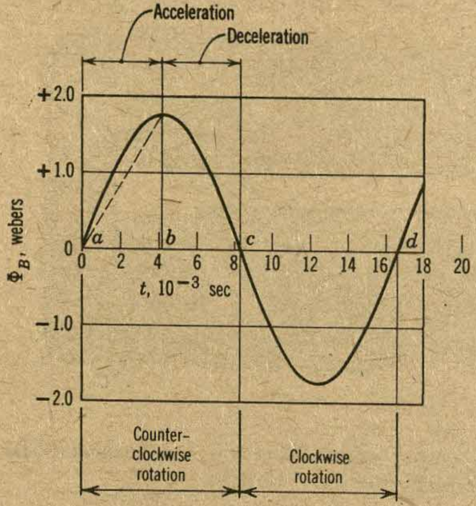


Fig. 35–15 The flux through the orbit of a betatron, during one cycle. Rotation of the electrons in the desired direction (counterclockwise as viewed from above in Fig. 35–14) is possible only during half-cycle *ac*.

From Faraday's law (Eq. 35–1), this is also the emf in volts. The electron will thus increase its energy by 430 ev every time it makes a trip around the orbit in the changing flux. If the electron gains only 430 ev of energy per revolution, it must make about 230,000 rev to gain its full 100 Mev. For an orbit radius of 33 in., this corresponds to a length of path of some 750 miles.

The betatron provides a good example of the fact that electric potential has no meaning for electric fields produced by induction. If a potential exists, it must be true that, as Eq. 35–10 shows, $\oint \mathbf{E}\cdot d\mathbf{l} = 0$ for any closed path. In the betatron, however, this integral, evaluated around the orbit, is precisely *not* zero but is, in our example, 430 volts. It must not be thought, of course, that the betatron violates the conservation of energy principle. The gain in kinetic energy of the circulating electron (430 ev/rev) must be supplied by an identifiable energy source. It comes, in fact, from the generator that energizes the magnet coils, thus providing the changing magnetic field. The energy is transmitted to the electron through the intermediary of this changing field.

▶ **Example 5.** In the betatron of Fig. 35–13, which is the "accelerating" quarter-cycle?

Let us assume that it is *ab* in Fig. 35–15, during which Φ_B through the orbit is *increasing*. If a conducting loop were placed to coincide with the orbit, an induced current would appear in the loop to oppose the tendency of Φ_B to increase. This means that a magnetic field would be set up that would oppose the field of the large magnet. Thus **E** would point outward at the right side of "doughnut" D in Fig. 35–14 and inward on the left side. The force ($-e\mathbf{E}$) acting on the electron is in the opposite direction to **E** because of the negative charge of the electron. Thus the tangential force acting on the electron is in the same direction as that at which it circulates in its orbit; this means that the speed of the electron will increase, as desired. The student should go through this same analysis carefully, assuming (incorrectly, as it will turn out) that the accelerating half-cycle is *bc* in Fig. 35–15 rather than *ab*.

35–7 Induction and Relative Motion

Faraday's law, in the form $\mathcal{E} = -d\Phi_B/dt$, gives correctly the induced emf no matter whether the change in Φ_B is produced by moving a coil, moving a magnet, changing the strength of a magnetic field, changing the shape of a conducting loop, or in other ways. However, observers who are in relative motion with respect to each other, even though they would all agree on the numerical value of the emf, would give different microscopic descriptions of the induction process. In electromagnetic systems, as well as in mechanical systems, it is important that the state of motion of the observer with respect to his environment be made perfectly clear.

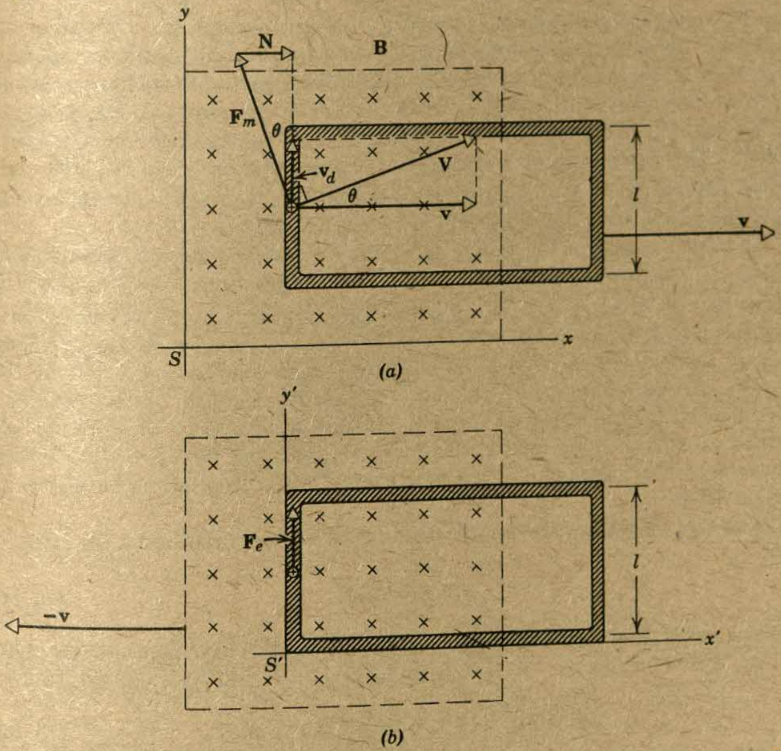


Fig. 35–16 A closed conducting loop in relative motion with respect to a magnetic field. (*a*) An observer *S*, fixed with respect to the magnet that produces the field **B**, sees the loop moving to the right. (*b*) An observer *S'*, fixed with respect to the loop, sees the magnet moving to the left.

Figure 35–16 shows a closed loop which is caused to move at velocity **v** with respect to a magnet that provides a uniform field **B** in the region shown. We consider first an observer, identified as *S*, who is *at rest with respect to the magnet* used to establish the field **B**; see Fig. 35–16*a*. The induced emf in this case is called a *motional emf* because the conducting loop is moving with respect to this observer.

Consider a positive charge carrier at the center of the left end of the conducting loop. To observer *S*, this charge, constrained to move to the right along with the loop, is a charge q moving, for an instant at least when the loop has just been set in motion, with a velocity $\mathbf{v}$ in the magnetic field $\mathbf{B}$ and as such it experiences a sideways magnetic deflecting force given by Eq. 33–3*a* ($\mathbf{F} = q\mathbf{v} \times \mathbf{B}$). This force causes the carriers to move upward along the conductor, so that they acquire a drift velocity $\mathbf{v}_d$ also, as shown in Fig. 35–16*a*.

The resultant equilibrium speed of the carrier when the velocity $\mathbf{v}$ of the loop has become constant is now $\mathbf{V}$, which we find by adding $\mathbf{v}$ and $\mathbf{v}_d$ vectorially in Fig. 35–16*a*. The magnetic deflecting force $\mathbf{F}_m$ is, as always, at right angles to the resultant velocity $\mathbf{V}$ of the carrier and is given by

$$\mathbf{F}_m = q\mathbf{V} \times \mathbf{B} \tag{35–13}$$

Now $\mathbf{F}_m$ acting alone would tend to push the carriers through the left wall of the conductor. Because this does not happen the conductor wall must exert a normal force $\mathbf{N}$ on the carriers (see Fig. 35–16*a*) of magnitude such that $\mathbf{v}_d$ lies parallel to the axis of the wire; in other words, $\mathbf{N}$ exactly cancels the horizontal component of $\mathbf{F}_m$, leaving only the component $\mathbf{F}_m \cos\theta$ that lies along the direction of the conductor. This component of force on the carrier is also cancelled out, this time by the average force $\overline{\mathbf{F}}_i$ associated with the internal collisions that the carrier experiences as it drifts with (constant) speed v_d through the wire.

The kinetic energy of the charge carrier as it drifts through the wire remains constant. This is consistent with the fact that the resultant force acting on the charge carrier ($= \mathbf{F}_m + \overline{\mathbf{F}}_i + \mathbf{N}$) is zero. The work done by $\mathbf{F}_m$ is zero because magnetic forces, ~~acting~~ at right angles to the velocity of a moving charge, can do no work on that charge. Thus the (negative) work done on the carrier by the average internal collision force $\mathbf{F}_i$ must be exactly cancelled by the (positive) work done on the carrier by the force $\mathbf{N}$. In the last analysis $\mathbf{N}$ is exerted by the agent who pulls the loop through the magnetic field, and the mechanical energy expended by this agent appears as heat energy in the loop, as we have seen in Section 35–4.

Let us then calculate the work dW done on the carrier in time dt by the force $\mathbf{N}$; it is

$$dW = N(v\,dt) \tag{35–14}$$

in which $v\,dt$ is the distance that the loop (and the carrier) have moved to the right in Fig. 35–16*a* in time dt. We can write for N (see Fig. 35–16*a* and Eq. 35–13).

$$N = F_m \sin\theta = (qVB)(v_d/V) = qBv_d \tag{35–15}$$

Substituting Eq. 35–15 into Eq. 35–14 yields

$$\begin{aligned} dW &= (qBv_d)(v\,dt) \\ &= (qBv)(v_d\,dt) = qBv\,dl \end{aligned} \tag{35–16}$$

in which $dl (= v_d\,dt)$ is the distance the carrier drifts along the conductor in time dt.

The work done on the carrier as it makes a complete circuit of the loop is found by integrating Eq. 35–16 around the loop and is

$$W = \oint dW = qBvl \tag{35-17}$$

This follows because work contributions for the top and the bottom of the loops are opposite in sign and cancel and no work is done in those portions of the loop that lie outside the magnetic field.

An agent that does work on charge carriers, thus establishing a current in a closed conducting loop, can be viewed as an emf. We can write, making use of Eq. 35-17,

$$\mathcal{E} = \frac{W}{q} = \frac{qBvl}{q} = Blv, \qquad (35\text{-}18)$$

which is, of course, the same result that we derived from Faraday's law of induction; see Eq. 35-3. Thus a motional emf is intimately connected with the sideways deflection of a charged particle moving through a magnetic field.

We now consider how the situation of Fig. 35-16 would appear to an observer S' who is *at rest with respect to the loop*. To this observer, the magnet is moving to the left in Fig. 35-16*b* with velocity $-\mathbf{v}$, and the charge q is at rest as far as its left-to-right motion is concerned. However S', like S, observes that the charge drifts clockwise around the loop and he measures the same emf $\mathcal{E}$ that S measures. S' accounts for this, at the microscopic level, by postulating that an electric field **E** is induced in the loop by the action of the moving magnet. This induced field **E**, which has the same origin as the induced fields that we discussed in Section 35-5, exerts a force on the charge carrier given by $q\mathbf{E}$.

The induced field **E**, which exists in the end of the loop only, is associated with an emf $\mathcal{E}$ and generates a current in the closed loop. Note that, in any closed loop in which there is a current, an internal electric field must exist at every point at which charges are moving. These electric fields however are set up by the emf, as in the case of a closed loop connected to a battery, and are *not* induced by the motion of the magnet. It is only this induced field **E** that we associate with the emf, through the relation (Eq. 35-8)

$$\mathcal{E} = \mathbf{E}\cdot d\mathbf{l}$$

which reduced to

$$\mathcal{E} = El \qquad (35\text{-}19)$$

in this case. This is so because no *induced* electric field appears in the upper and lower bars, because of the nature of their motions, and none appears in the part of the loop outside the magnetic field.

The emfs given by Eqs. 35-19 and 35-18 must be identical because the relative motion of the loop and the magnet is identical in the two cases shown in Fig. 35-16. Equating these relations yields

$$El = Blv,$$

or

$$E = vB. \qquad (35\text{-}20a)$$

In Fig. 35-16*b* the vector **E** points upward along the axis of the left end of the conducting loop because this is the direction in which positive charges are observed to drift. The directions of **v** and **B** are clearly shown in this figure. We see, then, that Eq. 35-20*a* is consistent with the more general vector relation

$$\mathbf{E} = \mathbf{v} \times \mathbf{B}. \qquad (35\text{-}20b)$$

We have not proved Eq. 35–20*b* except for the special case of Fig. 35–16; nevertheless it proves to be true in general, that is, no matter what the angle between **v** and **B**.

An experiment carried out in 1926 by the German physicist Wilhelm Wien (1864–1928) gives concrete support to our descriptions of the effect of the motion of an observer on the nature of the electric and magnetic fields that he observes. The radiations emitted from atoms vary slightly in wavelength if the atom is immersed in either a magnetic or an electric field. The atom can thus be used as a probe to examine the nature of such fields. Wien fired a beam of atoms with velocity **v** through a magnetic field **B**. The radiations emitted were identical in the distribution of their wavelengths with those that would have been emitted by a *resting* atom immersed in (*a*) the magnetic field **B** *and* (*b*) an induced electric field **E** given by Eq. 35–20*b*.

We interpret Eq. 35-20*b* in the following way: Observer S fixed with respect to the magnet is aware only of a magnetic field. The force to him arises from the motion of the charges through **B**. Observer S' fixed on the charge carrier is aware of an electric field **E** also and attributes the force on the charge (at rest with respect to him initially) to the electric field. S says the force is of purely magnetic origin and S' says the force is of purely electric origin. From the point of view of S, the induced emf is given by $\oint(\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{l}$. From the point of view of S', the same induced emf is given by $\oint \mathbf{E} \cdot d\mathbf{l}$, where **E** is the (induced) electric vector that he observes at points along the circuit.

For a third observer S'' who judges that both the magnet and the loop are moving, the force tending to move charges around the loop is neither purely electric nor purely magnetic, but a bit of each. In summary, in the equation

$$\mathbf{F}/q = \mathbf{E} + \mathbf{v} \times \mathbf{B},$$

different observers form different assessments of **E**, **B**, and **v** but, when these are combined, all observers form the same assessment of $\mathbf{F}/q$ and all obtain the same value for the induced emf in the loop (this depends only on the relative motion). That is, the total force is the same for all observers, but each observer forms a different estimate of the separate electric and magnetic forces contributing to the same total force.

▶ **Example 7.** In Fig. 35–16 assume that $B = 2.0$ webers/meter², $l = 10$ cm, and $v = 1.0$ meter/sec. Calculate (*a*) the induced electric field observed by S', and (*b*) the emf induced in the loop.

(*a*) The electric field, which is apparent only to observer S', is associated with the moving magnet and is given in magnitude (see Eq. 35–20*a*) by

$$\begin{aligned} E &= vB \\ &= (1.0 \text{ meter/sec})(2.0 \text{ webers/meter}^2) \\ &= 2.0 \text{ volt/meter.} \end{aligned}$$

(*b*) Observer S would calculate the induced (motional) emf from

$$\begin{aligned} \mathcal{E} &= Blv \\ &= (2.0 \text{ webers/meters}^2)(1.0 \times 10^{-1} \text{ meter})(1.0 \text{ meter/sec}) \\ &= 0.20 \text{ volt.} \end{aligned}$$

Observer S' would not regard the emf as motional and would use the relationship

$$\begin{aligned}\mathcal{E} &= El\\ &= (2.0\ \text{volt/meter})(1.0 \times 10^{-1}\ \text{meter})\\ &= 0.20\ \text{volt}.\end{aligned}$$

As must be the case, both observers agree as to the numerical value of the emf. ◀

QUESTIONS

1. The north pole of a magnet is moved away from a metallic ring, as in Fig. 35–17. In the part of the ring farthest from the reader, which way does the current point?

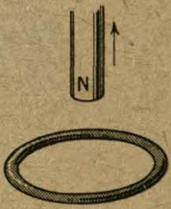


Fig. 35–17

2. *Eddy currents.* A sheet of copper is placed in a magnetic field as shown in Fig. 35–18. If we attempt to pull it out of the field or push it further in, an automatic resisting force appears. Explain its origin. (Hint: Currents, called eddy currents, are induced in the sheet in such a way as to oppose the motion.)

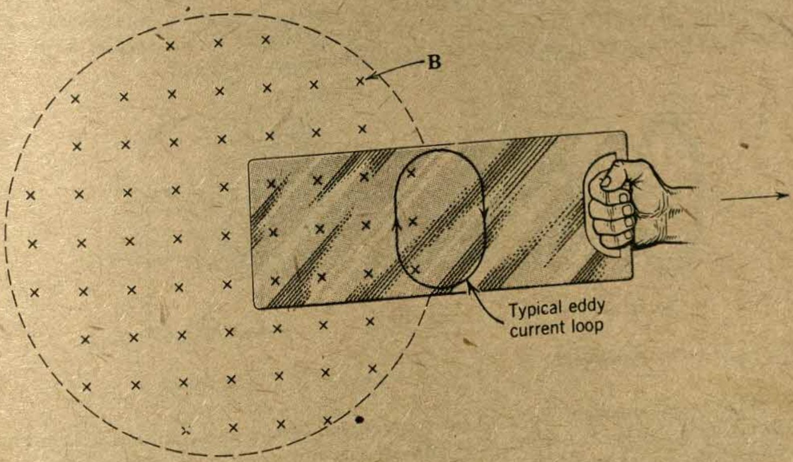


Fig. 35–18

3. *Electromagnetic shielding.* Consider a conducting sheet lying in a plane perpendicular to a magnetic field **B**, as shown in Fig. 35–19. (*a*) If **B** suddenly changes, the full *change* in **B** is not immediately detected in region *P*. Explain. (*b*) If the resistivity of the sheet is zero, the *change* is not ever detected at *P*. Explain. (*c*) If **B** changes periodically at high frequency and the conductor is made of a material of low resistivity, the region near *P* is almost completely shielded from the *changes* in flux. Explain. (*d*) Is such a conductor useful as a shield from *static* magnetic fields? Explain.

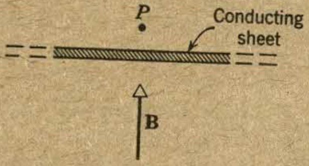


Fig. 35–19

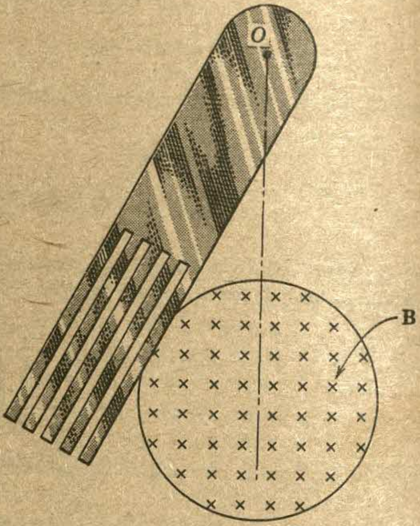


Fig. 35–20

4. *Magnetic damping.* A strip of copper is mounted as a pendulum about *O* in Fig. 35–20. It is free to swing through a magnetic field normal to the page. If the strip has slots cut in it as shown, it can swing freely through the field. If a strip without slots is substituted, the vibratory motion is strongly damped. Explain. (Hint: Use Lenz's law; consider the paths that the charge carriers in the strip must follow if they are to oppose the motion.)

5. Two conducting loops face each other a distance *d* apart (Fig. 35–21). An observer sights along their common axis. If a clockwise current *i* is suddenly established in the larger loop, (*a*) what is the direction of the induced current in the smaller loop? (*b*) What is the direction of the force (if any) that acts on the smaller loop?

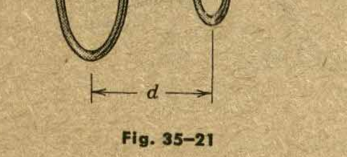


Fig. 35–21

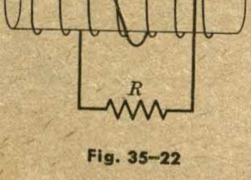


Fig. 35–22

6. What is the direction, if any, of the conventional current through resistor *R* in Fig. 35–22 (*a*) immediately after switch *S* is closed, (*b*) some time after switch *S* was closed, and (*c*) immediately after switch *S* is opened. (*d*) When switch *S* is held closed, which end of the coil acts as a north pole?

7. A current-carrying solenoid is moved toward a conducting loop as in Fig. 35-23. What is the direction of circulation of current in the loop as we sight toward it as shown?

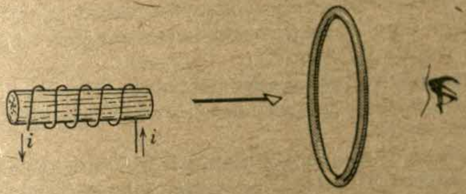


Fig. 35-23

8. If the resistance R in the left-hand circuit of Fig. 35-24 is increased, what is the direction of the induced current in the right-hand circuit?

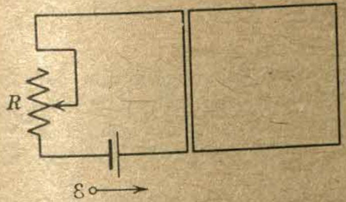


Fig. 35-24

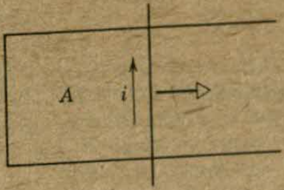


Fig. 35-25

9. In Fig. 35-25 the movable wire is moved to the right, causing an induced current as shown. What is the direction of $\mathbf{B}$ in region A?

10. A loop, shown in Fig. 35-26, is removed from the magnet by pulling it vertically upward. (*a*) What is the direction of the induced current? (*b*) Is a force required to remove the loop? (*c*) Does the total amount of Joule heat produced in removing the loop depend on the time taken to remove it?

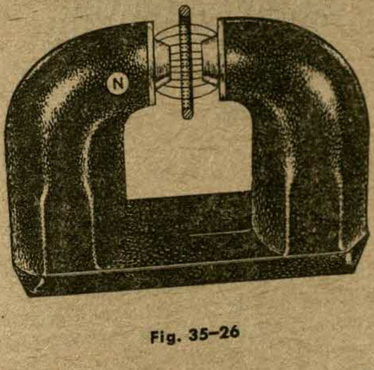


Fig. 35-26

11. A magnet is dropped down a long vertical copper tube. Show that, even neglecting air resistance, the magnet will reach a constant terminal velocity.

12. A magnet is dropped from the ceiling along the axis of a copper loop lying flat on the floor. If the falling magnet is photographed with a time sequence camera, what differences, if any, will be noted if (*a*) the loop is at room temperature and (*b*) the loop is packed in dry ice?

13. A copper ring and a wooden ring of the same dimensions are placed so that there is the same changing magnetic flux through each. How do the induced electric fields in each ring compare?

14. In Fig. 35–12 how can the induced emfs around paths 1 and 2 be identical? The induced electric fields are much weaker near path 1 than near path 2, as the spacing of the lines of force shows. See also Fig. 35–11.

15. In a certain betatron the electrons rotate counterclockwise as seen from above. In what direction must the magnetic field point and how must it change with time while the electron is being accelerated?

16. Why can a betatron be used for acceleration only during one-quarter of a cycle?

17. To make the electrons in a betatron orbit spiral outward, would it be necessary to increase or to decrease the central flux? Assume that **B** at the orbit remains essentially unchanged.

18. A cyclotron is a so-called *resonance device*. Does a betatron depend on resonance?

PROBLEMS

1. A hundred turns of insulated copper wire are wrapped around an iron cylinder of cross-sectional area 0.001 meter^2 and are connected to a resistor. The total resistance in the circuit is 10 ohms. If the longitudinal magnetic induction in the iron changes from 1 weber/meter^2 in one direction to 1 weber/meter^2 in the opposite direction, how much charge flows through the circuit?

2. In Fig. 35–27 a closed copper coil with 100 turns and a total resistance of 5.0 ohms is placed *outside* a solenoid like that of Example 1. If the current in the solenoid is changed as in that example, what current appears in the coil?

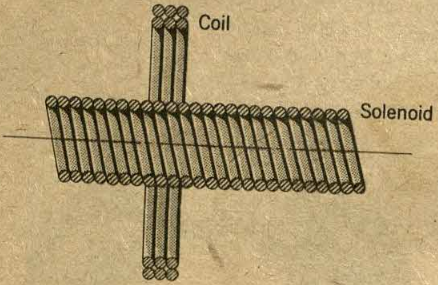


Fig. 35–27

3. A circular loop of wire 10 cm in diameter is placed with its normal making an angle of 30° with the direction of a uniform 5000-gauss magnetic field. The loop is "wobbled" so that its normal rotates about the field direction at the constant rate of 100 rev/min; the angle between the normal and the field direction (= 30°) remains unchanged during this process. What emf appears in the loop?

4. A uniform field of induction **B** is normal to the plane of a circular ring 10-cm in diameter made of #10 copper wire (diameter = 0.10 in.). At what rate must B change with time if an induced current of 10 amp is to appear in the ring?

5. A uniform field of induction **B** is changing in magnitude at a constant rate dB/dt. You are given a mass m of copper which is to be drawn into a wire of radius r and formed into a circular loop of radius R. Show that the induced current in the loop does not depend

on the size of the wire or of the loop and, assuming **B** perpendicular to the loop, is given by

$$i = \frac{m}{4\pi\rho\delta}\frac{dB}{dt},$$

where ρ is the resistivity and δ the density of copper.

6. You are given 50 cm of #18 copper wire (diameter = 0.040 in.). It is formed into a circular loop and placed at right angles to a uniform magnetic field that is increasing with time at the constant rate of 100 gauss/sec. At what rate is Joule heat generated in the loop?

7. A small bar magnet is pulled rapidly through a conducting loop, along its axis. Sketch qualitatively (*a*) the induced current and (*b*) the rate of Joule heating as a function of the position of the center of the magnet. Assume that the north pole of the magnet enters the loop first and that the magnet moves at constant speed. Plot the induced current as positive if it is clockwise as viewed along the path of the magnet.

8. *Alternating current generator*. A rectangular loop of N turns and of length a and width b is rotated at a frequency ν in a uniform field of induction **B**, as in Fig. 35–28. (*a*) Show that an induced emf given by

$$\mathcal{E} = 2\pi\nu NbaB \sin 2\pi\nu t = \mathcal{E}_0 \sin 2\pi\nu t$$

appears in the loop. This is the principle of the commercial alternating-current generator. (*b*) Design a loop that will produce an emf with $\mathcal{E}_0$ = 150 volts when rotated at 60 rev/sec in a field of magnetic induction of 5000-gauss.

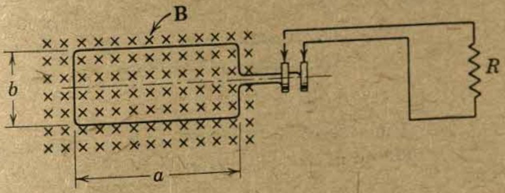


Fig. 35–28

9. A stiff wire bent into a semicircle of radius R is rotated with a frequency ν in a uniform field of induction **B**, as shown in Fig. 35–29. What are the amplitude and frequency of the induced voltage and of the induced current when the internal resistance of the meter M is R_M and the remainder of the circuit has negligible resistance?

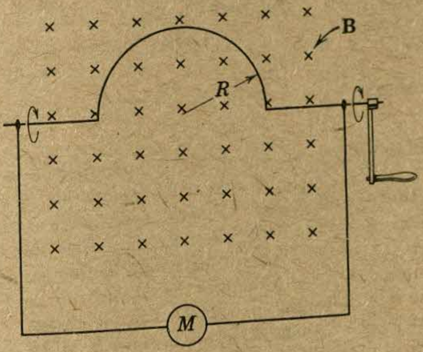


Fig. 35–29

10. A circular copper disk 10 cm in diameter rotates at 1800 rev/min about an axis through its center and at right angles to the disk. A uniform field of induction **B** of 10,000 gauss is perpendicular to the disk. What potential difference develops between the axis of the disk and its rim?

11. Figure 35–30 shows a copper rod moving with velocity **v** parallel to a long straight wire carrying a current i. Calculate the induced emf in the rod, assuming that $v = 5.0$ meters/sec, $i = 100$ amp, $a = 1.0$ cm, and $b = 20$ cm.

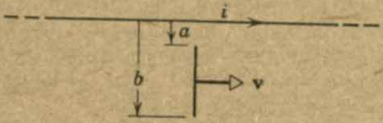


Fig. 35–30

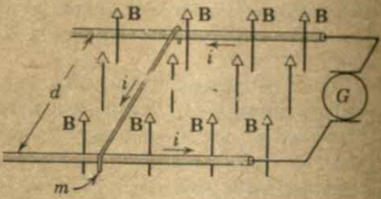


Fig. 35–31

12. A metal wire of mass m slides without friction on two rails spaced a distance d apart, as in Fig. 35–31. The track lies in a vertical uniform field of induction **B**. (*a*) A *constant current* i flows from generator G along one rail, across the wire, and back down the other rail. Find the velocity (speed and direction) of the wire as a function of time, assuming it to be at rest at $t = 0$. (*b*) The generator is replaced by a battery with *constant emf* $\mathcal{E}$. The velocity of the wire now approaches a constant final value. What is this terminal speed? (*c*) What is the current in part (*b*) when the terminal speed has been reached?

13. In Fig. 35–32 the magnetic flux through the loop perpendicular to the plane of the coil and directed into the paper is varying according to the relation

$$\Phi_B = 6t^2 + 7t + 1,$$

where Φ_B is in milliwebers (1 milliweber $= 10^{-3}$ weber) and t is in seconds. (*a*) What is the magnitude of the emf induced in the loop when $t = 2$ sec? (*b*) What is the direction of the current through R?

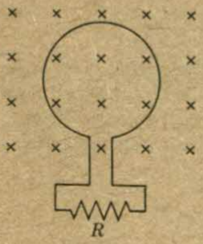


Fig. 35–32

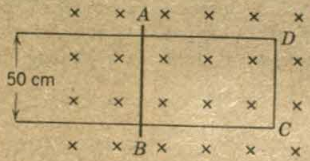


Fig. 35–33

14. In Fig. 35–33 a conducting rod AB makes contact with the metal rails AD and BC which are 50 cm apart in a uniform magnetic field of induction 1.0 weber/meter² perpendicular to the plane of the paper as shown. The total resistance of the circuit $ABCD$ is 0.4 ohm (assumed constant). (*a*) What is the magnitude and direction of the emf induced in the rod when it is moved to the left with a velocity of 8 meters/sec? (*b*) What force is required to keep the rod in motion? (*c*) Compare the rate at which mechanical work is done by the force **F** with the rate of development of heat in the circuit.

15. A square wire of length l, mass m, and resistance R slides without friction down parallel conducting rails of negligible resistance, as in Fig. 35–34. The rails are connected

to each other at the bottom by a resistanceless rail parallel to the wire, so that the wire and rails form a closed rectangular conducting loop. The plane of the rails makes an angle θ with the horizontal, and a uniform vertical field of magnetic induction **B** exists throughout the region. (*a*) Show that the wire acquires a steady-state velocity of magnitude

$$v = \frac{mgR \sin \theta}{B^2 l^2 \cos^2 \theta}.$$

(*b*) Prove that this result is consistent with the conservation-of-energy principle. (*c*) What change, if any, would there be if **B** were directed down instead of up?

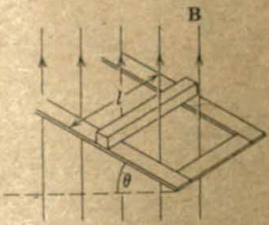


Fig. 35–34

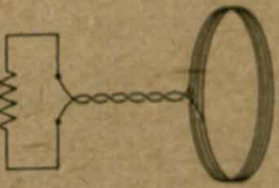


Fig. 35–35

16. Prove that if the flux of magnetic induction through the coil of N turns of Fig. 35–35 changes in any way from Φ_1 to Φ_2, then the charge q that flows through the circuit of total resistance R is given by

$$q = \frac{N(\Phi_2 - \Phi_1)}{R}.$$

17. Figure 35–36 shows a uniform field of induction **B** confined to a cylindrical volume of radius R. **B** is decreasing in magnitude at a constant rate of 100 gauss/sec. What is the instantaneous acceleration (direction and magnitude) experienced by an electron placed at a, at b, and at c? Assume $r = 5.0$ cm. (The necessary fringing of the field beyond R will not change your answer as long as there is axial symmetry about a perpendicular axis through b.)

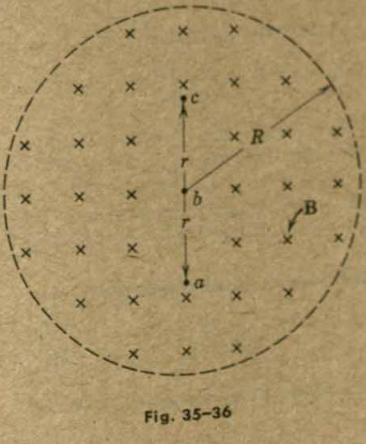


Fig. 35–36

18. A uniform magnetic field of induction **B** fills a cylindrical volume of radius R. A metal rod of length l is placed as shown in Fig. 35–37. If B is changing at the rate dB

show that the emf that is produced by the changing magnetic field and that acts between the ends of the rod is given by

$$\mathcal{E} = \frac{dB}{dt}\frac{l}{2}\sqrt{R^2 - \left(\frac{l}{2}\right)^2}.$$

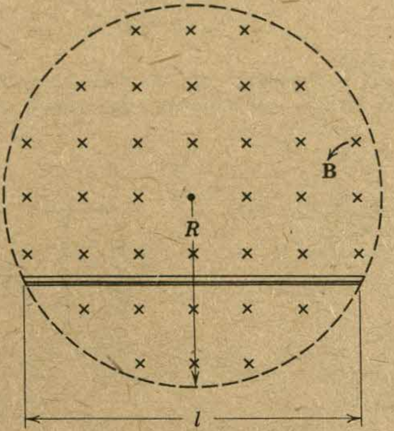


Fig. 35-37

19. Some measurements of the maximum magnetic induction as a function of radius for the General Electric Company betatron are as follows:

r, cm	B, gauss	r, cm	B, gauss
0	4000	81.2	4090
10.2	9500	83.7	4000
68.2	9500	88.9	3810
73.2	5280	91.4	3720
75.2	4510	93.5	3600
77.3	4280	95.5	3400

Show by graphical analysis that the relation $\bar{B} = 2B_R$ is satisfied at the orbit radius, $R = 84$ cm. (Hint: Note that $\bar{B} = \frac{1}{\pi R^2}\int_0^R B(r)(2\pi r)\,dr$ and evaluate the integral graphically.)

20. Prove that the electric field **E** in a charged parallel-plate capacitor cannot drop abruptly to zero as one moves at right angles to it, as suggested by the arrow in Fig. 35-38 (see point a). In actual capacitors fringing of the lines of force always occurs, which means that **E** approaches zero in a continuous and gradual way. See Problem 34-28. (Hint: Apply Faraday's law to the rectangular path shown by the dashed lines.)

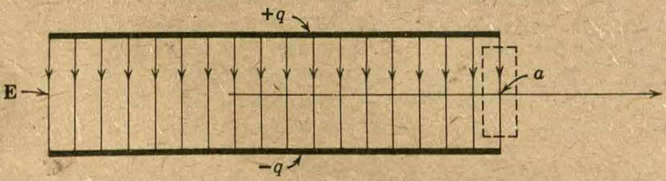


Fig. 35-38

Inductance

CHAPTER 36

36–1 Inductance

If two coils are near each other, a current i in one coil will set up a flux Φ_B through the second coil. If this flux is changed by changing the current, an induced emf will appear in the second coil according to Faraday's law. However, two coils are not needed to show an inductive effect. An induced emf appears in a coil if the current *in that same coil* is changed. This is called *self-induction* and the electromotive force produced is called a *self-induced emf.* It obeys Faraday's law of induction just as other induced emfs do.

Consider first a "close-packed" coil, a toroid, or the central section of a long solenoid. In all three cases the flux Φ_B set up in each turn by a current i is essentially the same for every turn. Faraday's law for such coils (Eq. 35–2)

$$\mathcal{E} = -\frac{d(N\Phi_B)}{dt} \tag{36–1}$$

shows that the number of *flux linkages* $N\Phi_B$ (N being the number of turns) is the important characteristic quantity for induction. For a given coil, provided no magnetic materials such as iron are nearby, this quantity is proportional to the current i, or

$$N\Phi_B = Li, \tag{36–2}$$

in which L, the proportionality constant, is called the *inductance* of the device.

From Faraday's law (see Eq. 36–1) the induced emf can be written as

$$\mathcal{E} = -\frac{d(N\Phi_B)}{dt} = -L\frac{di}{dt}. \tag{36–3a}$$

Written in the form

$$L = -\frac{\mathcal{E}}{di/dt}, \tag{36–3b}$$

this relation may be taken as the defining equation for inductance for coils of all shapes and sizes, whether or not they are close-packed and whether or not iron or other magnetic material is nearby. It is analogous to the defining relation for capacitance, namely

$$C = \frac{q}{V}.$$

If no iron or similar materials are nearby, L depends only on the geometry of the device. In an *inductor* (symbol) the presence of a *magnetic field* is the significant feature, corresponding to the presence of an *electric* field in a *capacitor*.

The unit of inductance, from Eq. 36–3b, is the volt-sec/amp. A special name, the *henry*, has been given to this combination of units, or

$$1 \text{ henry} = 1 \text{ volt-sec/amp}.$$

The unit of inductance is named after Joseph Henry (1797–1878), an American physicist and a contemporary of Faraday. Henry independently discovered the law of induction at about the same time Faraday did. The units *millihenry* (1 mh = 10^{-3} henry) and *microhenry* (1 μh = 10^{-6} henry) are also commonly used.

The direction of a self-induced emf can be found from Lenz's law. Suppose that a steady current i, produced by a battery, exists in a coil. Let us suddenly reduce the (battery) emf in the circuit to zero. The current i will start to decrease at once; this *decrease* in current, in the language of Lenz's law, is the "change" which the self-induction must oppose. To oppose the falling current, the induced emf must point in the same direction as the current, as in Fig. 36–1a. When the current in a coil is increased, Lenz's law shows that the self-induced emf points in the *opposite* direction to that of the current, as in Fig. 36–1b. In each case the self-induced emf acts to op-

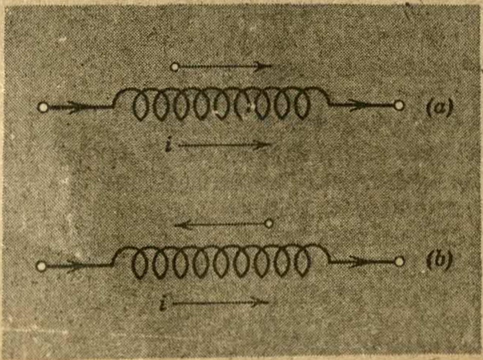


Fig. 36–1 In (a) the current i is *decreasing* and in (b) it is *increasing*. The self-induced emf $\mathcal{E}_L$ opposes the *change* in each case.

pose the *change* in the current. The minus sign in Eq. 36–3 shows that $\mathcal{E}$ and di/dt are opposite in sign, since L is always a positive quantity.

36–2 Calculation of Inductance

It has proved possible to make a direct calculation of capacitance in terms of geometrical factors for a few special cases, such as the parallel-plate capacitor. In the same way, it is possible to calculate the self-inductance L for a few special cases.

For a close-packed coil with no iron nearby, we have, from Eq. 36–2,

$$L = \frac{N\Phi_B}{i}. \tag{36–4}$$

Let us apply this equation to calculate L for a section of length l near the center of a long solenoid. The number of flux linkages in the length l of the solenoid is

$$N\Phi_B = (nl)(BA),$$

where n is the number of turns per unit length, B is the magnetic induction inside the solenoid, and A is the cross-sectional area. From Eq. 34–7, B is given by

$$B = \mu_0 ni.$$

Combining these equations gives

$$N\Phi_B = \mu_0 n^2 liA.$$

Finally, the inductance, from Eq. 36–4, is

$$L = \frac{N\Phi_B}{i} = \mu_0 n^2 lA. \tag{36–5}$$

The inductance of a length l of a solenoid is proportional to its volume (lA) and to the square of the number of turns per unit length. Note that it depends on geometrical factors only. The proportionality to n^2 is expected. If the number of turns per unit length is doubled, not only is the *total* number of turns N doubled but also the flux *through each turn* Φ_B is also doubled, an over-all factor of four for the flux linkages $N\Phi_B$, hence also a factor of four for the inductance (Eq. 36–4).

▶ **Example 1.** Derive an expression for the inductance of a toroid of rectangular cross section as shown in Fig. 36–2. Evaluate for $N = 10^3$, $a = 5.0$ cm, $b = 10$ cm, and $h = 1.0$ cm.

The lines of **B** for the toroid are concentric circles. Applying Ampère's law,

$$\oint \mathbf{B}\cdot d\mathbf{l} = \mu_0 i,$$

to a circular path of radius r yields

$$(B)(2\pi r) = \mu_0 i_0 N,$$

where N is the number of turns and i_0 is the current in the toroid windings; recall that

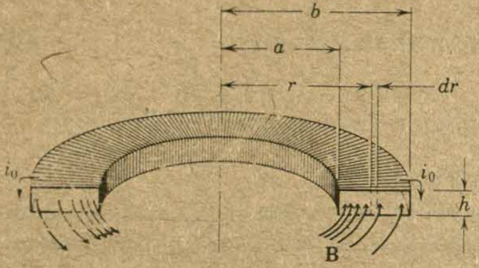


Fig. 36–2 Example 1. A cross section of a toroid, showing the current in the windings and the magnetic field.

i in Ampère's law is the *total* current that passes through the path of integration. Solving for B yields

$$B = \frac{\mu_0 i_0 N}{2\pi r}.$$

The flux Φ_B for the cross section of the toroid is

$$\Phi_B = \int \mathbf{B}\cdot d\mathbf{S} = \int_a^b (B)(h\,dr) = \int_a^b \frac{\mu_0 i_0 N}{2\pi r}\,h\,dr$$
$$= \frac{\mu_0 i_0 N h}{2\pi}\int_a^b \frac{dr}{r} = \frac{\mu_0 i_0 N h}{2\pi}\ln\frac{b}{a},$$

where $h\,dr$ is the area of the elementary strip shown in the figure.

The inductance follows from Eq. 36–4, or

$$L = \frac{N\Phi_B}{i_0} = \frac{\mu_0 N^2 h}{2\pi}\ln\frac{b}{a}.$$

Substituting numerical values yields

$$L = \frac{(4\pi \times 10^{-7}\ \text{weber/amp-m})(10^3)^2(1.0 \times 10^{-2}\ \text{meter})}{2\pi}\ln\frac{10 \times 10^{-2}\ \text{meter}}{5 \times 10^{-2}\ \text{meter}}$$
$$= 1.4 \times 10^{-3}\ \text{weber/amp} = 1.4\ \text{mh}.$$

36–3 An *LR* Circuit

In Section 32–8 we saw that if an emf $\mathcal{E}$ is suddenly introduced, perhaps by using a battery, into a single loop circuit containing a resistor R and a capacitor C the charge does not build up immediately to its final equilibrium value $(= C\mathcal{E})$ but approaches it in an exponential fashion described by Eq. 32–15, or

$$q = C\mathcal{E}\,(1 - e^{-t/\tau_c}). \tag{36–6}$$

The delay in the rise of the charge is described by the *capacitative time constant* τ_c, defined from

$$\tau_c = RC. \tag{36–7}$$

If in this same circuit the battery emf $\mathcal{E}$ is suddenly removed, the charge does not immediately fall to zero but approaches zero in an exponential fashion, described by Eq. 32–18b, or

$$q = C\mathcal{E}\,e^{-t/\tau_c}. \tag{36–8}$$

The same time constant τ_c describes the fall of the charge as well as its rise.

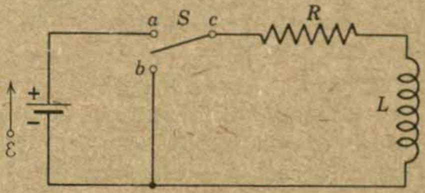


Fig. 36–3 An LR circuit

An analogous delay in the rise or fall of the current occurs if an emf $\mathcal{E}$ is suddenly introduced into, or removed from, a single loop circuit containing a resistor R and an inductor L. When the switch S in Fig. 36–3 is closed on a, for example, the current in the resistor starts to rise. If the inductor were not present, the current would rise rapidly to a steady value $\mathcal{E}/R$. Because of the inductor, however, a self-induced emf $\mathcal{E}_L$ appears in the circuit and, from Lenz's law, this emf opposes the *rise* of current, which means that it opposes the battery emf $\mathcal{E}$ in polarity. Thus the resistor responds to the difference between two emfs, a constant one $\mathcal{E}$ due to the battery and a variable one $\mathcal{E}_L$ $(= -L\,di/dt)$ due to self-induction. As long as this second emf is present, the current in the resistor will be less than $\mathcal{E}/R$.

As time goes on, the rate at which the current increases becomes less rapid and the self-induced emf $\mathcal{E}_L$, which is proportional to di/dt, becomes smaller. Thus a time delay is introduced, and the current in the circuit approaches the value $\mathcal{E}/R$ asymptotically.

When the switch S in Fig. 36–3 is thrown to a, the circuit reduces to that of Fig. 36–4. Let us apply the loop theorem, starting at x in this figure and going clockwise around the loop. For the direction of current shown, x will be higher in potential than y, which means that we encounter a drop in potential of $-iR$ as we traverse the resistor. Point y is higher in potential than point z because, for an increasing current, the induced emf will oppose the *rise* of the current by pointing as shown. Thus as we traverse the inductor from y to z we encounter a drop in potential of $-L(di/dt)$. We encounter a rise in potential of $+\mathcal{E}$ in traversing the battery from z to x. The loop theorem thus gives

$$-iR - L\frac{di}{dt} + \mathcal{E} = 0$$

or

$$L\frac{di}{dt} + iR = \mathcal{E}. \tag{36-9}$$

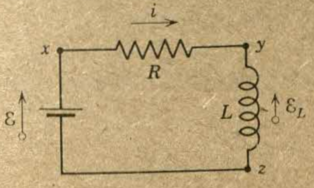


Fig. 36–4 The circuit of Fig. 36–3 just after switch S is closed on a.

Equation 36–9 is a *differential equation* involving the variable i and its first derivative di/dt. We seek the function $i(t)$ such that when it and its first derivative are substituted in Eq. 36–9 the equation is satisfied.

Although there are formal rules for solving various classes of differential equations (and Eq. 36–9 can, in fact, be easily solved by direct integration, after rearrangement) we often find it simpler to guess at the solution, guided by physical reasoning and by previous experience. Any proposed solution can be tested by substituting it in the differential equation and seeing whether this equation reduces to an identity.

The solution to Eq. 36–9 is, we assert,

$$i = \frac{\mathcal{E}}{R}(1 - e^{-Rt/L}). \tag{36–10}$$

To test this solution by substitution, we find the derivative di/dt, which is

$$\frac{di}{dt} = \frac{\mathcal{E}}{L}e^{-Rt/L}. \tag{36–11}$$

Substituting i and di/dt into Eq. 36–9 leads to an identity, as the student can easily verify. Thus Eq. 36–10 is a solution of Eq. 36–9. Figure 36–5 shows how the potential difference V_R across the resistor ($= iR$; see Eq. 36–10) and V_L across the inductor ($= L\,di/dt$; see Eq. 36–11) vary with time for particular values of $\mathcal{E}$, L, and R. The student should compare this figure carefully with the corresponding figure for an RC circuit (Fig. 32–11).

We can rewrite Eq. 36–10 as

$$i = \frac{\mathcal{E}}{R}(1 - e^{-t/\tau_L}), \tag{36–12}$$

in which τ_L, the *inductive time constant*, is given by

$$\tau_L = L/R. \tag{36–13}$$

The student should note the correspondence between Eqs. 36–12 and 36–6.

To show that the quantity τ_L ($= L/R$) has the dimensions of time, we put

$$\frac{1\text{ henry}}{\text{ohm}} = \frac{1\text{ }\cancel{\text{henry}}}{\cancel{\text{ohm}}}\left(\frac{1\text{ }\cancel{\text{volt}}\text{-sec}}{1\text{ }\cancel{\text{henry}}\text{-}\cancel{\text{amp}}}\right)\left(\frac{1\text{ }\cancel{\text{ohm}}\text{-}\cancel{\text{amp}}}{1\text{ }\cancel{\text{volt}}}\right) = 1\text{ sec.}$$

The first quantity in parenthesis is a conversion factor based on the defining equation for inductance [$L = -\mathcal{E}/(di/dt)$; Eq. 36–3*b*]. The second conversion factor is based on the relation $V = iR$.

The physical significance of the time constant follows from Eq. 36–12. If we put $t = \tau_L = L/R$ in this equation, it reduces to

$$i = \frac{\mathcal{E}}{R}(1 - e^{-1}) = (1 - 0.37)\frac{\mathcal{E}}{R} = 0.63\frac{\mathcal{E}}{R}.$$

Thus the time constant τ_L is that time at which the current in the circuit

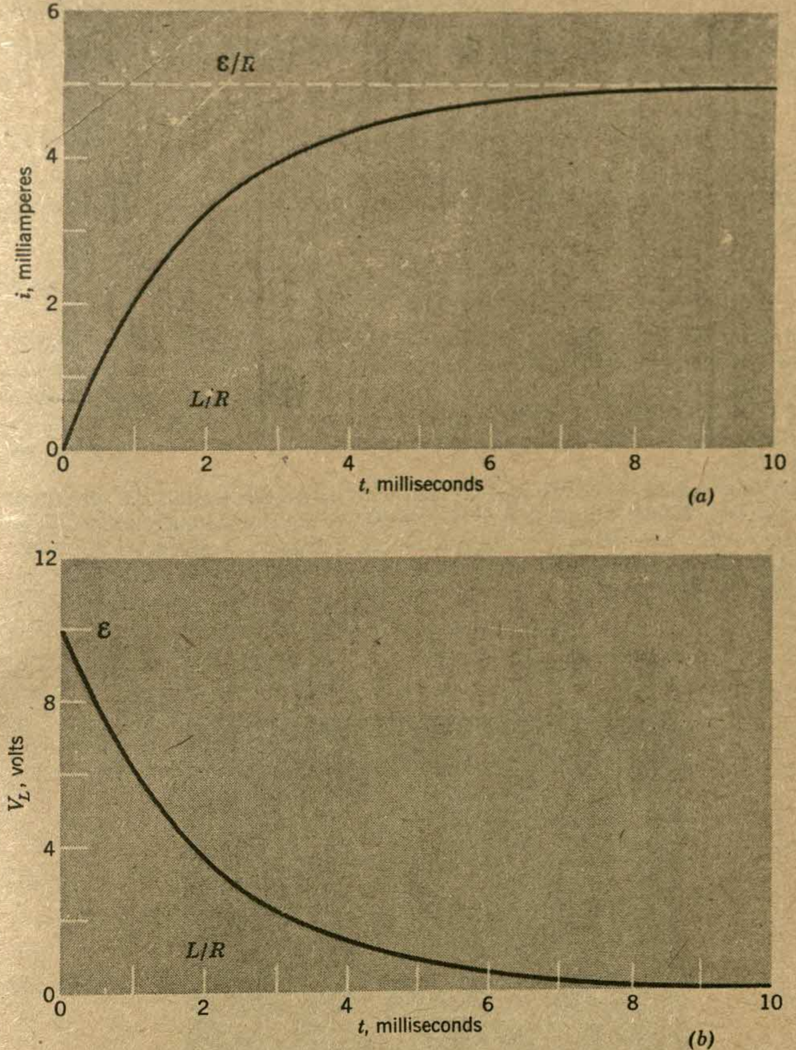


Fig. 36–5 If in Fig. 36–3 we assume that $R = 2000$ ohms, $L = 4$ henrys and $\mathcal{E} = 10$ volts, then (*a*) shows the variation of i with t during the current buildup after switch S is closed on a, and (*b*) the variations of V_L with t. The time constant is $L/R = 2.0 \times 10^{-3}$ sec.

will reach a value within $1/e$ (about 37%) of its final equilibrium value (see Fig. 36–5).

If the switch S in Fig. 36–3, having been left in position a long enough for the equilibrium current $\mathcal{E}/R$ to be established, is thrown to b, the effect is to remove the battery from the circuit. The differential equation that governs the subsequent decay of the current in the circuit can be found by putting $\mathcal{E} = 0$ in Eq. 36–9, or

$$L\frac{di}{dt} + iR = 0. \tag{36–14}$$

The student can show by the test of substitution that the solution of this differential equation is

$$i = \frac{\mathcal{E}}{R}e^{-t/\tau_L}. \tag{36–15}$$

Just as for the *RC* circuit, the behavior of the circuit of Fig. 36–3 can be investigated experimentally, using a cathode-ray oscilloscope. If switch *S* in this figure is thrown periodically between *a* and *b*, the applied emf alternates between the values $\mathcal{E}$ and zero. If the terminals of an oscilloscope are connected across *b* and *c* in Fig. 36–3, the oscilloscope will display the waveform of this applied emf on its screen, as in Fig. 36–6*c*.

If the terminals of the oscilloscope are connected across the resistor, the waveform displayed (Fig. 36–6*a*) will be that of the current in the circuit, since the potential drop across R, which determines the oscilloscope deflection, is given by $V_R = iR$. During the intervals marked *inc* in Fig. 36–6, the current is increasing and the waveform (see Eq. 36–12) is given by

$$V_R (= iR) = \mathcal{E}(1 - e^{-t/\tau_L}).$$

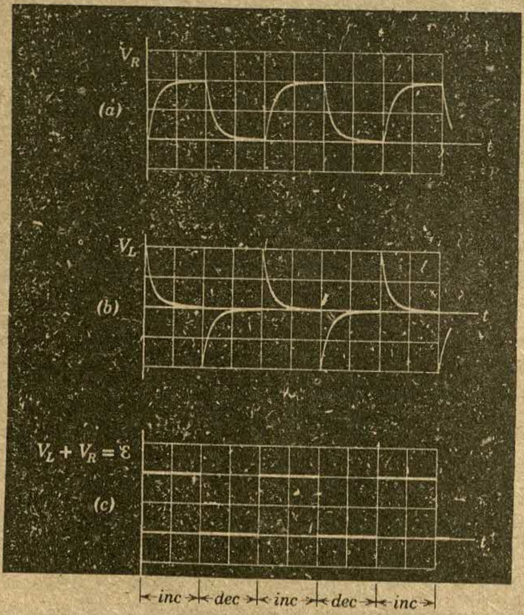


Fig. 36–6 Oscilloscope photograph showing the variation with time of (*a*) the potential drop V_R across the resistor, (*b*) the potential drop V_L across the inductor, and (*c*) the applied emf $\mathcal{E}$. During the intervals marked *inc* the current is increasing; during those marked *dec* it is decreasing. Compare with Fig. 32–13. (Courtesy E. K. Hege.)

During the intervals marked *dec*, the current is decreasing and V_R (see Eq. 36–15) is given by

$$V_R (= iR) = \mathcal{E}e^{-t/\tau_L}.$$

Note that both the growth and the decay of the current are delayed.

If the oscilloscope terminals are connected across the inductor, the screen will show a plot of the potential difference across it as a function of time (Fig. 36–6*b*). While the current is increasing, the equation of the trace (see Eq. 36–11) should be

$$V_L \left(= L\frac{di}{dt}\right) = \mathcal{E}e^{-t/\tau_L}.$$

When the current is decreasing, V_L is given in terms of the time derivative of Eq. 36–15 and is

$$V_L \left(= L\frac{di}{dt}\right) = -\mathcal{E}e^{-t/\tau_L}.$$

Note that V_L is opposite in sign when the current is increasing (di/dt positive) and when it is decreasing (di/dt negative), as is true also for the induced emf $\mathcal{E}_L[= -L(di/dt) = -V_L]$.

Examination of Fig. 36–6 shows that at any instant the sum of curves *a* and *b* always yields curve *c*. This is an expected consequence of the loop theorem, as Eq. 36–9 shows.

▶ **Example 2.** A solenoid has an inductance of 50 henrys and a resistance of 30 ohms. If it is connected to a 100-volt battery, how long will it take for the current to reach one-half its final equilibrium value?

The equilibrium value of the current is reached as $t \to \infty$; from Eq. 36–12 it is $\mathcal{E}/R$. If the current has half this value at a particular time t_0, this equation becomes

$$\frac{1}{2}\frac{\mathcal{E}}{R} = \frac{\mathcal{E}}{R}(1 - e^{-t_0/\tau_L}).$$

Solving for t_0 yields

$$t_0 = \tau_L \ln 2 = 0.69\frac{L}{R}.$$

Putting $\tau_L = L/R$ and using the values given, this reduces to

$$t_0 = 0.69\tau_L = 0.69\left(\frac{50 \text{ henrys}}{30 \text{ ohms}}\right) = 1.2 \text{ sec}.$$

36–4 Energy and the Magnetic Field

In Section 30–7 we saw that the electric field could be viewed as the site of stored energy, the energy per unit volume being given, in a vacuum, by

$$u_E = \tfrac{1}{2}\epsilon_0 E^2,$$

where E is the electric field strength at the point in question. Although this formula was derived for a parallel-plate capacitor, it holds for all kinds of electric field configurations.

Energy can also be stored in a magnetic field. For example, two parallel wires carrying currents in the same direction attract each other, and to pull

From Eq. 36-17, the desired rate is

$$P_B = \frac{dU_B}{dt} = Li\frac{di}{dt}$$

$$= (3.0 \text{ henrys})(0.189 \text{ amp})(0.37 \text{ amp/sec})$$

$$= 0.210 \text{ watt}.$$

Note that as required by the principle of conservation of energy (see Eq. 36-16)

$$P_{\mathcal{E}} = P_J + P_B,$$

or

$$0.567 \text{ watt} = 0.357 \text{ watt} + 0.210 \text{ watt}$$

$$= 0.567 \text{ watt}.$$

◀

36-5 Energy Density and the Magnetic Field

We now derive an expression for the *density* of energy u in a magnetic field. Consider a length l near the center of a very long solenoid; Al is the volume associated with this length. The stored energy must lie entirely within this volume because the magnetic field outside such a solenoid is essentially zero. Moreover, the stored energy must be uniformly distributed throughout the volume of the solenoid because the magnetic field is uniform everywhere inside.

Thus we can write

$$u_B = \frac{U_B}{Al}$$

or, since

$$U_B = \tfrac{1}{2}Li^2,$$

$$u_B = \frac{\tfrac{1}{2}Li^2}{Al}.$$

To express this in terms of the magnetic field, we can substitute for L in this equation, using the relation $L = \mu_0 n^2 lA$ (Eq. 36-5). Also we can solve Eq. 34-7 ($B = \mu_0 in$) for i and substitute in this equation. Doing so yields finally

$$u_B = \frac{1}{2}\frac{B^2}{\mu_0}. \qquad (36\text{-}19)$$

This equation gives the energy density stored at any point (in a vacuum or in a nonmagnetic substance) where the magnetic induction is **B**. The equation is true for all magnetic field configurations, even though it was derived by considering a special case, the solenoid. Equation 36-19 is to be compared with Eq. 30-27,

$$u_E = \tfrac{1}{2}\epsilon_0 E^2, \qquad (\kappa = 1) \qquad (36\text{-}20)$$

which gives the energy density (in a vacuum) at any point in an electric field. Note that both u_B and u_E are proportional to the square of the appropriate field quantity, B or E.

The solenoid plays a role with relationship to magnetic fields similar to the role the parallel-plate capacitor plays with respect to electric fields. In each case we have a simple device that can be used for setting up a uniform field throughout a well-defined region of space and for deducing, in a simple way, some properties of these fields.

▶ **Example 5.** A long *coaxial cable* (Fig. 36-7) consists of two concentric cylinders with radii a and b. Its central conductor carries a steady current i, the outer conductor providing the return path. (a) Calculate the energy stored in the magnetic field for a length l of such a cable.

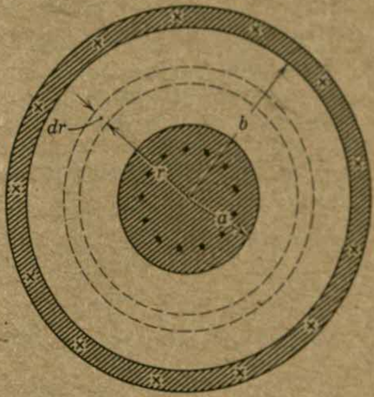


Fig. 36-7 Example 5. Cross section of a coaxial cable, showing steady currents in the central and outer conductors.

In the space between the two conductors Ampère's law

$$\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0 i$$

leads to

$$(B)(2\pi r) = \mu_0 i$$

or

$$B = \frac{\mu_0 i}{2\pi r}.$$

Ampère's law shows further that the magnetic field is zero for points outside the outer conductor (why?). Magnetic fields exist *inside* each of the conductors; although their values can readily be found from Ampère's law, we choose to ignore them, on the assumption that the cable dimensions are chosen so that most of the stored magnetic energy is in the space between the conductors.

The energy density for points between the conductors, from Eq. 36-19, is

$$u = \frac{1}{2\mu_0} B^2 = \frac{1}{2\mu_0} \left(\frac{\mu_0 i}{2\pi r}\right)^2 = \frac{\mu_0 i^2}{8\pi^2 r^2}.$$

Consider a volume element dV consisting of a cylindrical shell whose radii are r and $r + dr$ and whose length is l. The energy dU contained in it is

$$dU = u\, dV = \frac{\mu_0 i^2}{8\pi^2 r^2} (2\pi r l)(dr) = \frac{\mu_0 i^2 l}{4\pi} \frac{dr}{r}.$$

The total stored magnetic energy is found by integration, or

$$U = \int dU = \frac{\mu_0 i^2 l}{4\pi} \int_a^b \frac{dr}{r} = \frac{\mu_0 i^2 l}{4\pi} \ln \frac{b}{a},$$

which is the desired expression.

(*b*) What is the inductance of a length l of coaxial cable?

The inductance L can be found from Eq. 36–18 ($U = \frac{1}{2}Li^2$), which leads to

$$L = \frac{2U}{i^2} = \frac{\mu_0 l}{2\pi} \ln \frac{b}{a}.$$

The student should also derive this expression directly from the definition of inductance, using the procedures of Example 1.

Example 6. Compare the energy required to set up, in a cube 10 cm on edge, (*a*) a uniform electric field of 10^5 volts/meter and (*b*) a uniform magnetic field of 1 weber/meter^2 (= 10^4 gauss). Both these fields would be judged reasonably large but they are readily available in the laboratory.

(*a*) In the electric case we have, where V_0 is the volume of the cube,

$$\begin{aligned} U_E &= u_E V_0 = \tfrac{1}{2}\epsilon_0 E^2 V_0 \\ &= (0.5)(8.9 \times 10^{-12}\ \text{coul}^2/\text{nt-m}^2)(10^5\ \text{volts/meter})^2(0.1\ \text{meter})^3 \\ &= 4.5 \times 10^{-5}\ \text{joule}. \end{aligned}$$

(*b*) In the magnetic case, from Eq. 36–19, we have

$$\begin{aligned} U_B &= u_B V_0 = \frac{B^2}{2\mu_0} V_0 = \frac{(1\ \text{weber/meter}^2)^2(0.1\ \text{meter})^3}{(2)(4\pi \times 10^{-7}\ \text{weber/amp-m})} \\ &= 400\ \text{joules}. \end{aligned}$$

In terms of fields normally available in the laboratory, much larger amounts of energy can be stored in a magnetic field than in an electric one, the ratio being about 10^7 in this example. Conversely, much more energy is required to set up a magnetic field of reasonable laboratory magnitude than is required to set up an electric field of similarly reasonable magnitude. ◀

QUESTIONS

1. Two coils are connected in series. Does their equivalent inductance depend on their geometrical relationship to each other?

2. Is the inductance per unit length for a solenoid near its center (*a*) the same as, (*b*) less than, or (*c*) greater than the inductance per unit length near its ends?

3. Two solenoids, A and B, have the same diameter and length and contain only one layer of windings, with adjacent turns touching, insulation thickness being negligible. Solenoid A contains many turns of fine wire and solenoid B contains fewer turns of heavier wire. (*a*) Which solenoid has the larger inductance? (*b*) Which solenoid has the larger inductive time constant?

4. If the flux passing through each turn of a coil is the same, the inductance of the coil may be computed from $L = N\Phi_B/i$ (Eq. 36–4). How might one compute L for a coil for which this assumption is not valid?

5. If a current in a source of emf is in the direction of the emf, the energy of the source decreases; if a current is in a direction opposite to the emf (as in charging a battery), the energy of the source increases. Do these statements apply to the inductor in Fig. 36–1*a* and 36–1*b*?

6. Show that the dimensions of the two expressions for L, $N\Phi_B/i$ (Eq. 36–4) and $\mathcal{E}/(di/dt)$ (Eq. 36–3*b*), are the same.

7. You are given N turns of wire connected in series. How should the turns be arranged to obtain the maximum self-inductance?

8. Does the time required for the current in a particular LR circuit to build up to any given fraction of its equilibrium value depend on the value of the applied emf?

9. A steady current is set up in a coil with a very large inductive time constant. When the current is interrupted with a switch, a heavy arc tends to appear at the switch blades. Explain. (Note: Interrupting currents in highly inductive circuits can be dangerous.)

10. In an LR circuit like that of Fig. 36–4 can the self-induced emf ever be larger than the battery emf?

11. In an LR circuit like that of Fig. 36–4 is the current in the resistance *always* the same as the current in the inductance?

12. In the circuit of Fig. 36–3 the self-induced emf is a maximum at the instant the switch is closed on a. How can this be since there is no current in the inductance at this instant?

13. Give some arguments to show that energy can be stored in a magnetic field.

14. The switch in Fig. 36–3 is thrown from a to b. What happens to the energy stored in the inductor?

15. In a toroid is the energy density larger near the inner radius or near the outer radius?

PROBLEMS

1. A 10-henry inductor carries a steady current of 2.0 amp. How can a 100-volt self-induced emf be made to appear in the inductor?

2. Two inductances L_1 and L_2 are connected in series and are separated by a large distance. (*a*) Show that the equivalent inductance L is $L_1 + L_2$. (*b*) Why must their separation be large?

3. Show that if two inductors with equal inductance L are connected in parallel the equivalent inductance of the combination is $\frac{1}{2}L$. The inductors are separated by a large distance.

4. Two parallel wires whose centers are a distance d apart carry equal currents in opposite directions. Show that, neglecting the flux within the wires themselves, the inductance of a length l of such a pair of wires is given by

$$L = \frac{\mu_0 l}{\pi} \ln \frac{d - a}{a},$$

where a is the wire radius. See Example 4, Chapter 34.

5. A long thin solenoid can be bent into a ring to form a toroid. Show that if the solenoid is long and thin enough the equation for the inductance of a toroid (see Example 1) reduces to that for a solenoid (Eq. 36–5).

6. A solenoid is wound with a single layer of #10 copper wire (diameter, 0.10 in.). It is 4.0 cm in diameter and 2.0 meters long. What is the inductance per unit length for the solenoid near its center? Assume that adjacent wires touch and that insulation thickness is negligible.

7. The inductance of a close-packed coil of 400 turns is 8 mh. What is the magnetic flux through the coil when the current is 5×10^{-3} amp?

8. A wooden toroidal core with a square cross section has an inner radius of 10 cm and an outer radius of 12 cm. It is wound with one layer of #18 wire (diameter, 0.040 in.; resistance, 160 ft/ohm). What are (*a*) the inductance and (*b*) the inductive time constant? Ignore the thickness of the insulation.

9. The current in an LR circuit builds up to one-third of its steady-state value in 5.0 sec. What is the inductive time constant?

10. How many "time constants" must we wait for the current in an LR circuit to build up to within 0.1 per cent of its equilibrium value?

11. The switch S in Fig. 36–3 is thrown from b to a. After one inductive time const show that (*a*) the total energy transformed to Joule heat in the resistor is $0.168\mathcal{E}^2\tau$

and that (*b*) the energy stored in the magnetic field is $0.200\mathcal{E}^2\tau_L/R$. (*c*) Show that the equilibrium energy stored in the magnetic field is $0.500\mathcal{E}^2\tau_L/R$.

12. Show that the inductive time constant τ_L can also be defined as the time that would be required for the current in an LR circuit to reach its equilibrium value *if it continued to increase at its initial rate.*

13. A 50-volt potential difference is suddenly applied to a coil with $L = 50$ mh and $R = 180$ ohms. At what rate is the current increasing after 0.001 sec?

14. A coil with an inductance of 2.0 henrys and a resistance of 10 ohms is suddenly connected to a resistanceless battery with $\mathcal{E} = 100$ volts. At 0.1 sec after the connection is made, what are the rates at which (*a*) energy is being stored in the magnetic field, (*b*) Joule heat is appearing, and (*c*) energy is being delivered by the battery?

15. A coil with an inductance of 2.0 henrys and a resistance of 10 ohms is suddenly connected to a resistanceless battery with $\mathcal{E} = 100$ volts. (*a*) What is the equilibrium current? (*b*) How much energy is stored in the magnetic field when this current exists in the coil?

16. Prove that when switch S in Fig. 36–3 is thrown from a to b all the energy stored in the inductor appears as Joule heat in the resistor.

17. A circular loop of wire 5.0 cm in radius carries a current of 100 amp. What is the energy density at the center of the loop?

18. What is the magnetic energy density at the center of a circulating electron in the hydrogen atom (see Example 9, Chapter 34)?

19. A long wire carries a current of uniform density. Let i be the total current carried by the wire and show that the magnetic energy per unit length stored *within* the wire equals $\mu_0 i^2/16\pi$. Note that it does not depend on the wire diameter.

20. Show that the self-inductance for a length l of a long wire associated with the flux *inside* the wire only is $\mu_0 l/8\pi$, independent of the wire diameter.

21. The coaxial cable of Example 5 has $a = 1.0$ mm, $b = 4.0$ mm, and $c = 5.0$ mm (c is the radius of the outer surface of the outer conductor). It carries a current of 10 amp in the inner conductor and an equal but oppositely directed return current in the outer conductor. Calculate and compare the stored magnetic energy per meter of cable length (*a*) within the central conductor, (*b*) in the space between the conductors, and (*c*) within the outer conductor.

22. A length of #10 copper wire carries a current of 10 amp. Calculate (*a*) the magnetic energy density and (*b*) the electric energy density at the surface of the wire. The wire diameter is 0.10 in. and its resistance per unit length is 1.0 ohm/1000 ft.

23. What must be the strength of a uniform electric field if it is to have the same energy density as that possessed by a 5000-gauss magnetic field?

Magnetic Properties of Matter

CHAPTER 37

37-1 Poles and Dipoles

In electricity the *isolated charge* q is the simplest structure that can exist. If two such charges of opposite sign are placed near each other, they form an *electric dipole*, characterized by an electric dipole moment $\mathbf{p}$. In magnetism isolated magnetic "poles," which would correspond to isolated electric charges, apparently do not exist. The simplest magnetic structure is the *magnetic dipole*, characterized by a magnetic dipole moment $\boldsymbol{\mu}$. Table 34-1 summarizes some characteristics of electric and magnetic dipoles.

A current loop, a bar magnet, and a solenoid of finite length are examples of magnetic dipoles. Their magnetic dipole moments can be measured by placing the dipole in an external magnetic field $\mathbf{B}$, measuring the torque $\boldsymbol{\tau}$ that acts on it, and computing $\boldsymbol{\mu}$ from Eq. 33-11, or

$$\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B}. \tag{37-1}$$

Alternatively, we can measure $\mathbf{B}$ due to the dipole at a point along its axis a (large) distance r from its center and compute μ from the expression in Table 34-1, or

$$B = \frac{\mu_0}{2\pi} \frac{\mu}{r^3}. \tag{37-2}$$

Figure 37-1, which shows iron filings sprinkled on a sheet of paper under which there is a bar magnet, suggests that this dipole might be viewed as two "poles" separated by a distance d. However, all attempts to isolate these poles fail. If the magnet is broken, as in Fig. 37-2, the fragments prove to be dipoles and not isolated poles. If we break up a magnet into the electrons

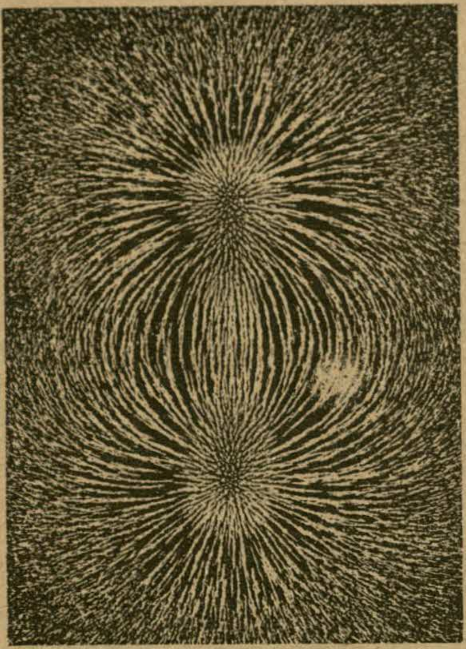


Fig. 37–1 A bar magnet is a magnetic dipole. The iron filings suggest the pattern of lines of force in Fig. 37–4*a*. (Courtesy Physical Science Study Committee.)

and nuclei that make up its atoms, it will be found that even these elementary particles are magnetic dipoles. Figure 37–3 contrasts the electric and the magnetic characters of the free electron.

All electrons have a characteristic *"spin" angular momentum* about a certain axis, which has the value of

$$L_s = 0.52723 \times 10^{-34} \text{ joule-sec.}$$

This is suggested by the vector $\mathbf{L}_s$ in Fig. 37–3*b*. Such a spinning charge can be viewed classically as being made up of infinitesimal current loops. Each such loop is a tiny magnetic dipole, its moment being given by (Eq. 33–10)

$$\mu = NiA, \tag{37-3}$$

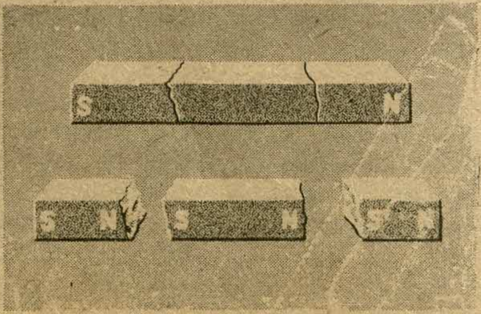
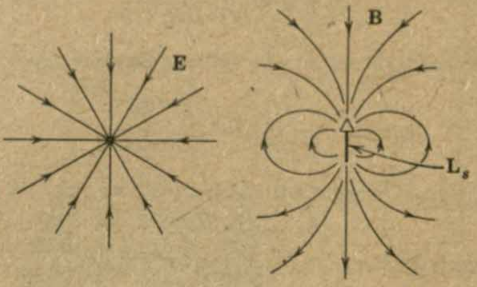


Fig. 37–2 If a bar magnet is broken, each fragment becomes a small dipole.

where i is the equivalent current in each infinitesimal loop and A is the loop area. The number of turns, N, is unity for each loop. The magnetic dipole moment of the spinning charge can be found by integrating over the moments of the infinitesimal current loops that make it up; see Problem 2.

Although this model of the spinning electron is too mechanistic and is not in accord with modern quantum physics, it remains true that the magnetic dipole moments of elementary particles are closely connected with their intrinsic angular momenta, or spins. Those particles and nuclei whose spin angular momentum is zero (the α-particle, the pion, the O^{16} nucleus, etc.) have no magnetic dipole moment. The "intrinsic" or "spin" magnetic moment of the electron must be distinguished from any additional magnetic moment it may have because of its *orbital* motion in an atom; see Example 2.



(a) (b)

Fig. 37-3 (*a*) The lines of **E** and (*b*) the lines of **B** for an electron. The magnetic dipole moment of the electron, μ_l, is directed opposite to the spin angular momentum vector, $\mathbf{L}_s$.

▶ **Example 1.** Devise a method for measuring μ for a bar magnet.

(*a*) Place the magnet in a uniform external magnetic field **B**, with $\boldsymbol{\mu}$ making an angle θ with **B**. The magnitude of the torque acting on the magnet (see Eq. 37-1) is given by

$$\tau = \mu B \sin \theta.$$

Clearly μ can be learned if τ, B, and θ are measured.

(*b*) A second technique is to suspend the magnet from its center of mass and to allow it to oscillate about its stable equilibrium position in the external field **B**. For *small* oscillations, $\sin \theta$ can be replaced by θ and the equation just given becomes

$$\tau = -(\mu B)\theta = -\kappa\theta,$$

where κ is a constant. The minus sign has been inserted to show that τ is a *restoring torque*. Since τ is proportional to θ, the condition for simple angular harmonic motion is met. The frequency ν is given by the reciprocal of Eq. 15-24, or

$$\nu = \frac{1}{2\pi}\sqrt{\frac{\kappa}{I}} = \frac{1}{2\pi}\sqrt{\frac{\mu B}{I}}.$$

With this equation μ can be found from the measured quantities ν, B, and I.

Example 2. An electron in an atom circulating in an assumed circular orbit of radius r behaves like a tiny current loop and has an *orbital magnetic dipole moment*[*] usually represented by μ_l. Derive a connection between μ_l and the *orbital angular momentum* L_l.

Newton's second law ($F = ma$) yields, if we substitute Coulomb's law for F,

$$\frac{1}{4\pi\epsilon_0}\frac{e^2}{r^2} = ma = \frac{mv^2}{r}$$

or

$$v = \sqrt{\frac{e^2}{4\pi\epsilon_0 mr}}. \qquad (37\text{–}4)$$

The angular velocity ω is given by

$$\omega = \frac{v}{r} = \sqrt{\frac{e^2}{4\pi\epsilon_0 mr^3}}.$$

The current for the orbit is the rate at which charge passes any given point, or

$$i = e\nu = e\left(\frac{\omega}{2\pi}\right) = \sqrt{\frac{e^4}{16\pi^3\epsilon_0 mr^3}}.$$

The orbital dipole moment μ_l is given from Eq. 37–3 if we put $N = 1$ and $A = \pi r^2$, or

$$\mu_l = NiA = (1)\sqrt{\frac{e^4}{16\pi^3\epsilon_0 mr^3}}\,(\pi r^2) = \frac{e^2}{4}\sqrt{\frac{r}{\pi\epsilon_0 m}}. \qquad (37\text{–}5)$$

The orbital angular momentum L_l is

$$L_l = (mv)r.$$

Combining with Eq. 37–4 leads to

$$L_l = \sqrt{\frac{e^2 mr}{4\pi\epsilon_0}}.$$

Finally, eliminating r between this equation and Eq. 37–5 yields

$$\mu_l = L_l\left(\frac{e}{2m}\right),$$

which shows that the orbital magnetic moment of an electron is proportional to its orbital angular momentum.

For $r = 5.1 \times 10^{-11}$ meter, which corresponds to hydrogen in its normal state, we have from Eq. 37–5

$$\begin{aligned}\mu_l &= \frac{e^2}{4}\sqrt{\frac{r}{\pi\epsilon_0 m}}\\ &= \frac{(1.6 \times 10^{-19}\ \text{coul})^2}{4}\sqrt{\frac{5.1 \times 10^{-11}\ \text{meter}}{(\pi)(8.9 \times 10^{-12}\ \text{coul}^2/\text{nt-m}^2)(9.1 \times 10^{-31}\ \text{kg})}}\\ &= 9.1 \times 10^{-24}\ \text{amp-m}^2.\end{aligned}$$

◀

---

[*] This must not be confused with the magnetic dipole moment μ_s of the electron spin, which is also present.

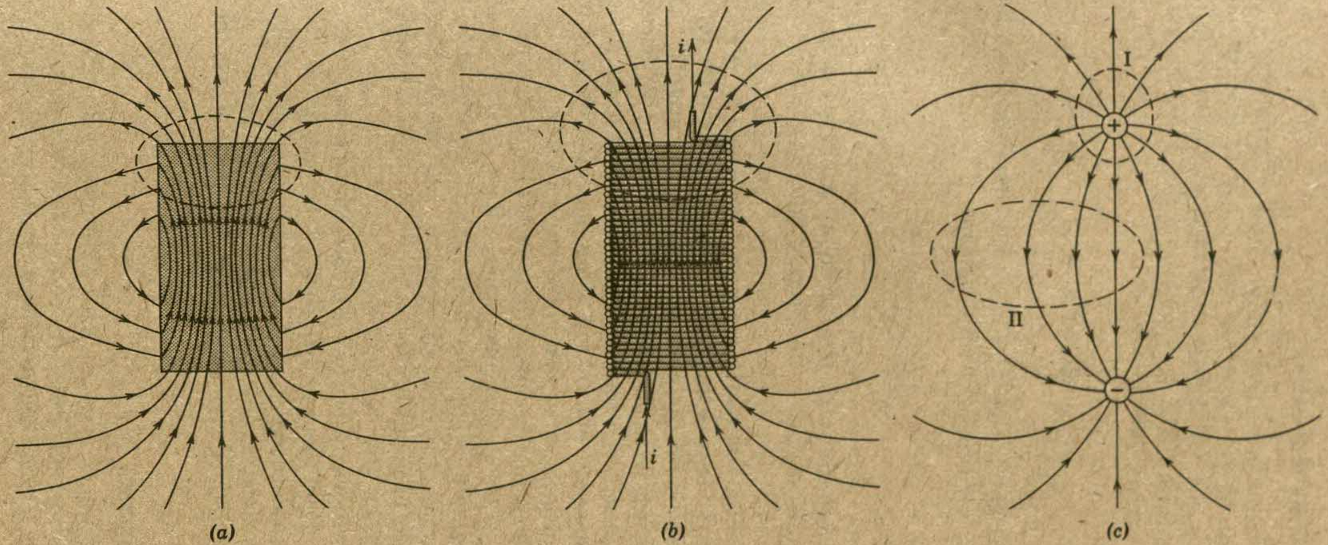


Fig. 37–4 Lines of **B** (*a*) for a bar magnet and (*b*) for a short solenoid. (*c*) Lines of **E** for an electric dipole. At large enough distances all three fields vary like those for a dipole. The four dashed curves are intersections with the plane of the figure of closed Gaussian surfaces. Note that Φ_B equals zero for (*a*) or (*b*). Φ_E equals zero for surfaces like II in (*c*), which do not contain any charge, but Φ_E is not zero for surfaces like I.

37–2 Gauss's Law for Magnetism

Gauss's law for magnetism, which is one of the basic equations of electromagnetism (see Table 38–3), is a formal way of stating a conclusion that seems to be forced on one by the facts of magnetism, namely, that *isolated magnetic poles do not exist.* This equation asserts that the flux Φ_B through any *closed* Gaussian surface must be zero, or

$$\Phi_B = \oint \mathbf{B}\cdot d\mathbf{S} = 0, \tag{37–6}$$

where the integral is to be taken over the entire closed surface. We contrast this with Gauss's law for electricity, which is

$$\epsilon_0 \oint \mathbf{E}\cdot d\mathbf{S} = q. \tag{37–7}$$

The fact that a zero appears at the right of Eq. 37–6, but not at the right of Eq. 37–7, means that in magnetism there is no counterpart to the free charge q in electricity.

Figure 37–4*a* shows a Gaussian surface that encloses one end of a bar magnet. Note that the lines of **B** enter the surface for the most part inside the magnet and leave it for the most part outside the magnet. There is thus an inward (or negative) flux inside the magnet and an outward (or positive) flux outside it. The total flux for the whole surface is zero. In Section 37–7 the interested student will find a fuller discussion of why the lines of **B** for a bar magnet are drawn in the way shown.

Figure 37–4*b* shows a similar surface for a solenoid of finite length which, like a bar magnet, is also a magnetic dipole. Here, too, Φ_B equals zero. Figures 37–4*a* and *b* show clearly that there are no "sources" of **B**; that is, there are no points from which lines of **B** emanate. Also, there are no "sinks" of **B**; that is, there are no points toward which **B** converges. In other words, *there are no free magnetic poles.*

Figure 37–4*c* shows a Gaussian surface (I) surrounding the positive end of an electric dipole. Here there *is* a net flux of the lines of **E**. There is a "source" of **E**; it is the charge q. If q is negative, we have a "sink" of **E** because the lines of **E** end on negative charges. For surfaces like surface II in Fig. 37–4*c* for which the charge inside is zero, the flux of **E** over the surface is also zero.

37–3 Paramagnetism

Magnetism as we know it in our daily experience is an important but special branch of the subject called *ferromagnetism;* we discuss this in Section 37–5. Here we discuss a weaker form of magnetism called *paramagnetism.*

For most atoms and ions, the magnetic effects of the electrons, including both their spins and orbital motions, exactly cancel so that the atom or ion is not magnetic. This is true for the rare gases such as neon and for ions* such as Cu^+, which makes up ordinary copper. These materials do not exhibit paramagnetism. For other atoms or ions the magnetic effects of the electrons do not cancel, so that the atom as a whole has a magnetic dipole moment $\boldsymbol{\mu}$. Examples are found among the so-called transition elements, such as Mn^{++}, the rare earths, such as Gd^{+++}, and the actinide elements, such as U^{++++}.

If a sample of N atoms, each of which has a magnetic dipole moment $\boldsymbol{\mu}$, is placed in a magnetic field, the elementary atomic dipoles tend to line up with the field. For perfect alignment, the sample as a whole would have a magnetic dipole moment of $N\boldsymbol{\mu}$. However, the aligning process is seriously interfered with by the collisions that take place between the atoms if the sample is a gas and by temperature vibrations if the sample is a solid. The importance of this thermal agitation effect may be measured by comparing two energies: one ($= \frac{3}{2}kT$) is the mean kinetic energy of translation of a gas atom at temperature T; the other ($= 2\mu B$) is the difference in energy between an atom lined up with the magnetic field and one pointing in the opposite direction. As Example 3 shows, the effect of the collisions at ordinary temperatures and fields is very great. The sample acquires a magnetic moment when placed in an external magnetic field, but this moment is usually much smaller than the maximum possible moment $N\mu$.

▶ **Example 3.** A paramagnetic gas, whose atoms (see Example 2) have a magnetic dipole moment of about 10^{-23} amp-m², is placed in an external magnetic field of magnitude 1 weber/meter². At room temperature ($T = 300°K$) calculate and compare U_T, the mean kinetic energy of translation ($= \frac{3}{2}kT$), and U_B, the magnetic energy ($= 2\mu B$):

$$U_T = \tfrac{3}{2}kT = (\tfrac{3}{2})(1.38 \times 10^{-23}\ \text{joule}/°\text{K})(300°\text{K}) = 6 \times 10^{-21}\ \text{joule},$$

$$U_B = 2\mu B = (2)(10^{-23}\ \text{amp-m}^2)(1\ \text{weber/meter}^2) = 2 \times 10^{-23}\ \text{joule}.$$

Because U_T equals 300 U_B, we see that energy exchanges in collisions can interfere seriously with the alignment of the dipoles with the external field. ◀

If a specimen of a paramagnetic substance is placed in a nonuniform magnetic field, such as that near the pole of a strong magnet, it will be attracted toward the region of higher field, that is, toward the pole. We can understand this by drawing an analogy with the corresponding electric case of Fig. 37–5, which shows a dielectric specimen (a sphere) in a nonuniform electric field. The net electric force points to the right in the figure and is

$$F_e = q(E + \Delta E) - q(E - \Delta E) = q(2\Delta E),$$

which can be written as

$$F_e = \frac{(q\,\Delta x)}{\Delta x}\,2\Delta E = p\left(\frac{2\Delta E}{\Delta x}\right) \cong p\left(\frac{dE}{dx}\right)_{\text{max}}.$$

Here $p\ (= q\,\Delta x)$ is the induced electric dipole moment of the sphere. In the differential limit of a very small sphere $(2\Delta E/\Delta x)$ approaches $(dE/dx)_{\text{max}}$, the gradient of the electric field at the center of the sphere.

---

[*] Cu^+ indicates a copper atom from which one electron has been removed; Al^{+++} indicates an aluminum atom from which three electrons have been removed, etc.

In the corresponding magnetic case we have, by analogy,

$$F_m = \mu \left(\frac{dB}{dx}\right)_{\max} \tag{37-8}$$

Thus, by measuring the magnetic force F_m that acts on a small paramagnetic specimen when it is placed in a nonuniform magnetic field whose field gradient $(dB/dx)_{\max}$ is known, we can learn its magnetic dipole moment μ. The ***magnetization*** $\mathbf{M}$ is defined as the magnetic moment per unit volume, or

$$\mathbf{M} = \frac{\boldsymbol{\mu}}{V},$$

where V is the volume of the specimen. It is a vector because $\boldsymbol{\mu}$, the dipole moment of the specimen, is a vector.

In 1895 Pierre Curie (1859–1906) discovered experimentally that the magnetization M of a paramagnetic specimen is directly proportional to B, the effective value of magnetic induction in which the specimen is placed, and inversely proportional to the temperature, or

$$M = C\frac{B}{T}, \tag{37-9}$$

in which C is a constant. This equation is known as *Curie's law.* The law is physically reasonable in that increasing B tends to align the elementary dipoles in the specimen, that is, to increase M, whereas increasing T tends to interfere with this alignment, that is, to decrease M. Curie's law is well verified experimentally, provided that the ratio B/T does not become too large.

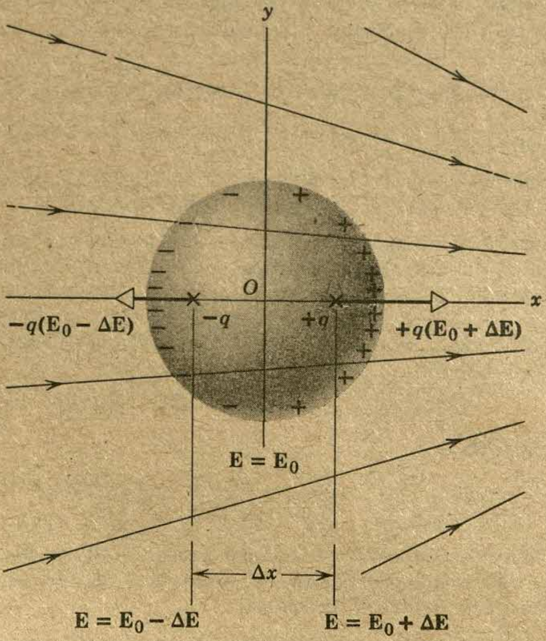

y
O
x
$-q(\mathbf{E}_0 - \Delta\mathbf{E})$
$-q$
$+q$
$+q(\mathbf{E}_0 + \Delta\mathbf{E})$
$\mathbf{E} = \mathbf{E}_0$
Δx
$\mathbf{E} = \mathbf{E}_0 - \Delta\mathbf{E}$
$\mathbf{E} = \mathbf{E}_0 + \Delta\mathbf{E}$

Fig. 37-5 A dielectric sphere in a nonuniform electric field.

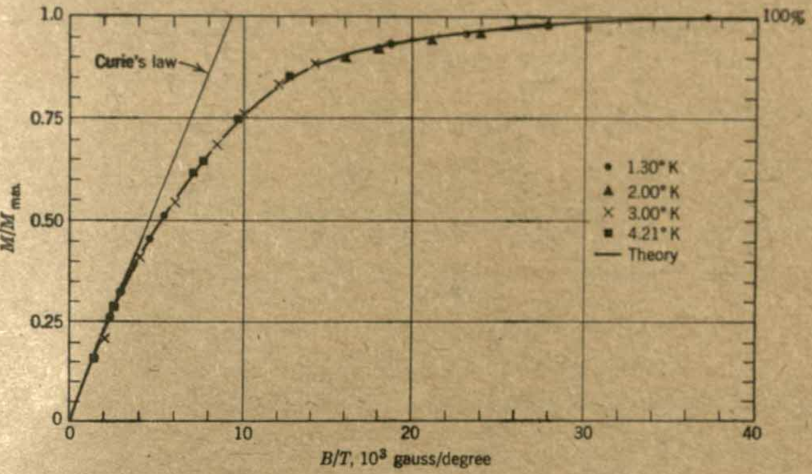


Fig. 37-6 The ratio $M/M_{\max}$ for a paramagnetic salt (chromium potassium alum) in various magnetic fields and at various temperatures. The curve through the experimental points is a theoretical curve calculated from modern quantum physics. (From measurements by W. E. Henry.)

M cannot increase without limit, as Curie's law implies, but must approach a value $M_{\max}$ $(= \mu N/V)$ corresponding to the complete alignment of the N dipoles contained in the volume V of the specimen. Figure 37-6 shows this saturation effect for a sample of $CrK(SO_4)_2 \cdot 12H_2O$. The chromium ions are responsible for all the paramagnetism of this salt, all the other elements being paramagnetically inert. To achieve 99.5% saturation, it is necessary to use applied magnetic fields as high as 50,000 gauss and temperatures as low as 1.3°K. Note that for more readily achievable conditions, such as $B = 10{,}000$ gauss and $T = 10°\mathrm{K}$, the abscissa in Fig. 37-6 is only 1.0 so that Curie's law would appear to be well obeyed for this and for all lower values of B/T. The curve that passes through the experimental points in this figure is calculated from a theory based on modern quantum physics; it is in excellent agreement with experiment.

37-4 Diamagnetism

In 1846 Michael Faraday discovered that a specimen of bismuth brought near to the pole of a strong magnet is *repelled*. He called such substances *diamagnetic*. Diamagnetism, present in all substances, is such a feeble effect that its presence is masked in substances made of atoms that have a net magnetic dipole moment, that is, in paramagnetic or ferromagnetic substances.

Figures 37-7*a* and *b* show an electron circulating in a diamagnetic atom at angular frequency ω_0 in an assumed circular orbit of radius r. Each electron is moving under the action of a centripetal force $\mathbf{F}_E$ of electrostatic origin where, from Newton's second law,

$$F_E = ma = m\omega_0^2 r. \tag{37-10}$$

Each rotating electron has an orbital magnetic moment, but for the atom as a whole the orbits are randomly oriented so that there is no *net* magnetic effect. In Fig. 37-7*a*, for example, the magnetic dipole moment $\boldsymbol{\mu}_l$ points into the page; in Fig.

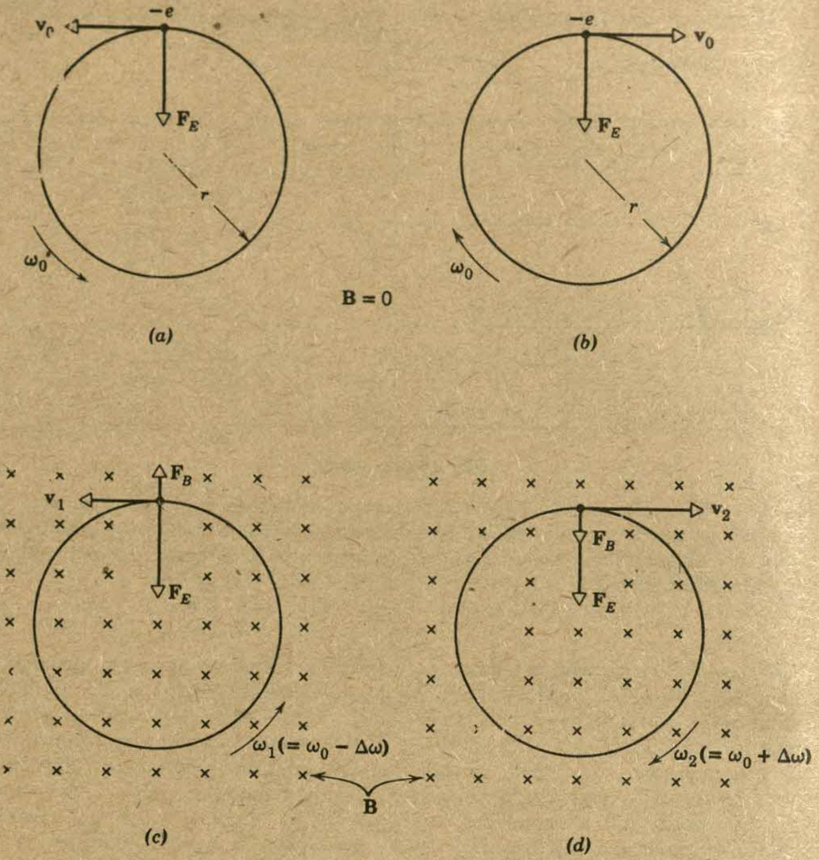


Fig. 37–7 (*a*) An electron circulating in an atom. (*b*) An electron circulating in the opposite direction. (*c*) A magnetic field is introduced, *decreasing* the linear speed of the electron in (*a*), that is, $v_1 < v_0$. (*d*) The magnetic field *increases* the linear speed of the electron in (*b*), that is, $v_2 > v_0$.

37–7*b* it points out and the net effect for the two orbits shown is cancellation. This cancellation is also shown at the left in Fig. 37–8.

If an external field **B** is applied as in Fig. 37–7*c* and *d*, an *additional* force, given by $-e(\mathbf{v} \times \mathbf{B})$, acts on the electron. This magnetic force acts always at right angles to the direction of motion; its magnitude is

$$F_B = evB = e(\omega r)B. \tag{37–11}$$

The student should show that in Fig. 37–7*c* $\mathbf{F}_B$ and $\mathbf{F}_E$ point in opposite directions and that in Fig. 37–7*d* they point in the same direction. Note that since the centripetal force changes when the magnetic field is applied (the radius can be shown to remain constant), the angular velocity must also change; thus ω in Eq. 37–11 differs from ω_0 in Eq. 37–10.

Applying Newton's second law to Figs. 37–7*c* and *d*, and allowing for both directions of circulation, yields for the *resultant* forces on the electrons

$$F_E \pm F_B = ma = m\omega^2 r.$$

Substituting Eqs. 37-10 and 37-11 into this equation yields

$$m\omega_0^2 r \pm e\omega r B = m\omega^2 r$$

or

$$\omega^2 \mp \left(\frac{eB}{m}\right)\omega - \omega_0^2 = 0. \tag{37-12}$$

This quadratic equation can be solved for ω, the new angular velocity. Rather than doing this, we take advantage of the fact (presented without proof; see Problem 7) that ω differs only slightly from ω_0, even in the strongest external magnetic fields. Thus

$$\omega = \omega_0 + \Delta\omega \tag{37-13}$$

where $\Delta\omega \ll \omega_0$. Substituting this equation into Eq. 37-12 yields

$$[\omega_0^2 + 2\omega_0\,\Delta\omega + (\Delta\omega)^2] \pm [\beta\omega_0 + \beta\,\Delta\omega] - \omega_0^2 = 0,$$

where β is a convenient abbreviation for eB/m. The two terms ω_0^2 cancel each other; the terms $(\Delta\omega)^2$ and $\beta\,\Delta\omega$ are small compared to the remaining terms and may be set equal to zero with only small error. This leads, as an excellent approximation, to

$$\Delta\omega \cong \mp \tfrac{1}{2}\beta = \mp \frac{eB}{2m} \tag{37-14}$$

or, from Eq. 37-13,

$$\omega = \omega_0 \mp \frac{eB}{2m}.$$

Thus the effect of applying a magnetic field is to increase or decrease (depending on the direction of circulation) the angular velocity. This, in turn, increases or decreases the orbital magnetic moment of the circulating electron (see Example 2).

In Fig. 37-7*c* the angular velocity is reduced (because the centripetal force is reduced) so that the magnitude of the magnetic moment is reduced. In Fig. 37-7*d*, however, the angular velocity is increased so that the magnitude of $\boldsymbol{\mu}_l$ is increased. These effects are shown on the right in Fig. 37-8, where it will be noted that the two magnetic moments *no longer cancel*.

We see that if a magnetic field **B** is applied to a diamagnetic substance (zero net magnetic moment in absence of applied field), a magnetic moment will be *induced* whose direction (out of the plane of Fig. 37-7) is *opposite* to **B**; see also Fig. 37-8. This is precisely the reverse of paramagnetism, in which the (*permanent*) magnetic dipoles tend to point in the *same* direction as the applied field.

We can now understand why a diamagnetic specimen is repelled when brought near to the pole of a strong magnet. If the pole is a north pole, there exists a nonuniform magnetic field of induction with **B** pointing away from the pole. If a sphere

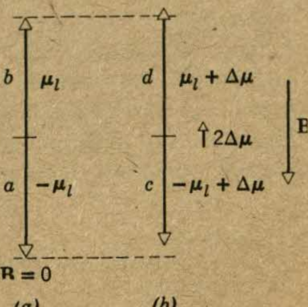


Fig. 37-8 The magnetic moments of the two oppositely circulating electrons in an atom cancel when there is no external magnetic field, as in (*a*), but do *not* cancel when a field is applied, as in (*b*). Note that the resultant moment in (*b*) points in the *opposite* direction to **B**. Compare carefully with Fig. 37-7.

made of a diamagnetic material (bismuth, say) is brought near to this pole, the magnetization **M** that is induced in it points toward the pole, that is, *opposite to* **B**. Thus the side of the sphere closest to the magnet behaves like a north pole and is *repelled* by the nearby north pole of the magnet. For a paramagnetic sphere, the vector **M** points *along the direction of* **B** and the side of the sphere nearest to the magnet is a south pole, which is *attracted* to the north pole of the magnet.

▶ **Example 4.** Calculate the *change* in magnetic moment for a circulating electron, as described in Example 2, if a magnetic field of induction **B** of 2.0 webers/meter² acts at right angles to the plane of the orbit.

We obtain μ from Eq. 37–3, or

$$\mu = NiA = (1)(e\nu)(\pi r^2) = (1)\left(\frac{e\omega}{2\pi}\right)(\pi r^2) = \tfrac{1}{2}er^2\omega.$$

The *change* in μ is

$$\Delta\mu = \tfrac{1}{2}er^2\,\Delta\omega$$

or, from Eq. 37–14,

$$\Delta\mu = \pm\,\tfrac{1}{2}er^2\left(\frac{eB}{2m}\right) = \pm\,\frac{e^2Br^2}{4m}.$$

Substituting numbers yields

$$\Delta\mu = \pm\,\frac{(1.6\times 10^{-19}\text{ coul})^2(2.0\text{ webers/meter}^2)(5.1\times 10^{-11}\text{ meter})^2}{(4)(9.1\times 10^{-31}\text{ kg})}$$

$$= \pm 3.7\times 10^{-29}\text{ amp-m}^2.$$

In Example 2 the moment μ_l was 9.1×10^{-24} amp-m², so that the change induced by even a strong external magnetic field is rather small, the ratio $\Delta\mu/\mu_l$ being about 4×10^{-6}. ◀

37–5 Ferromagnetism

For five elements (Fe, Co, Ni, Gd, and Dy) and for a variety of alloys of these and other elements a special effect occurs which permits a specimen to achieve a high degree of magnetic alignment in spite of the randomizing tendency of the thermal motions of the atoms. In such materials, described as *ferromagnetic*, a special form of interaction called *exchange coupling* occurs between adjacent atoms, coupling their magnetic moments together in rigid parallelism.* Modern quantum physics successfully predicts that this will occur only for the five elements listed. If the temperature is raised above a certain critical value, called the *Curie temperature*, the exchange coupling suddenly disappears and the materials become simply paramagnetic. For iron the Curie temperature is 1043°K. Ferromagnetism is evidently a property not only of the individual atom or ion but also of the interaction of each atom or ion with its neighbors in the crystal lattice (see Fig. 21–5) of the solid.

Figure 37–9 shows a *magnetization curve* for a specimen of iron. To obtain such a curve, we form the specimen, assumed initially unmagnetized, into a ring and wind a toroidal coil around it as in Fig. 37–10, to form a so-called *Rowland ring*. When a current i is set up in the coil, *if the iron core is not present*, a field of induction is set up within the toroid given by (Eq. 34–4)

$$B_0 = \mu_0 ni, \tag{37–15}$$

---

* Exchange coupling, a purely quantum effect, cannot be "explained" in terms of classical physics.

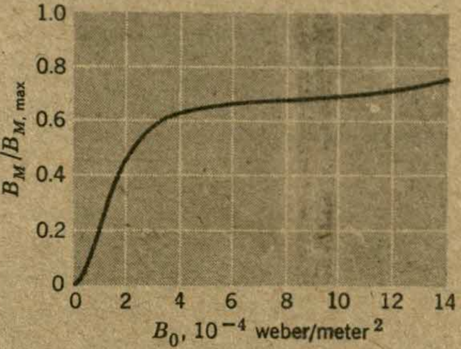


Fig. 37-9 A magnetization curve for iron.

where n is the number of turns per unit length for the toroid. Although this formula was derived for a long solenoid, it can be applied to a toroid if $d \ll r$ in Fig. 37-10. Because of the iron core, the actual value of **B** in the toroidal space will exceed $\mathbf{B}_0$, by a large factor in many cases, since the elementary atomic dipoles in the core line up with the applied field $\mathbf{B}_0$, thereby setting up their own field of induction. Thus we can write

$$B = B_0 + B_M \tag{37-16}$$

where B_M is the magnetic induction due to the specimen; it is proportional to the magnetization M of the specimen. Often $B_M \gg B_0$.

The field B_0 is proportional to the current in the toroid and can be calculated readily, using Eq. 37-15; B can be measured in a way that is described below. An experimental value for B_M can be derived from Eq.

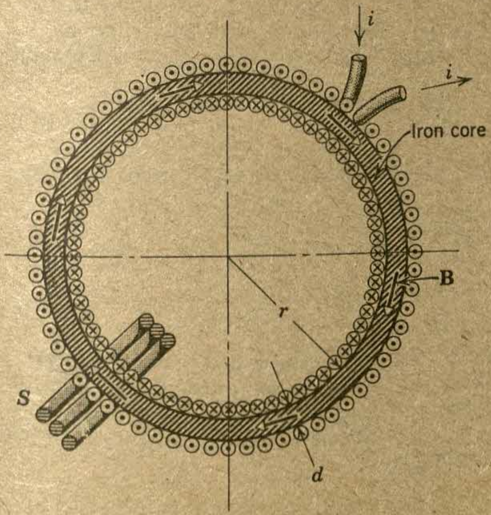


Fig. 37-10 A Rowland ring, showing a secondary coil S.

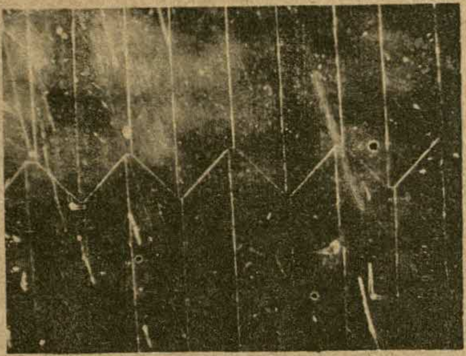


Fig. 37–14 Domain patterns for a single crystal of iron containing 3.8% silicon. The white lines show the boundaries between the domains. These boundaries are regular rather than irregular, as in Fig. 37–13, because the specimen is a single crystal. In Fig. 37–13 the specimen is made up of many crystallites or grains. (Courtesy H. J. Williams, Bell Telephone Laboratories.)

Figure 37–14 shows some domain photographs, taken by sprinkling a colloidal suspension of finely powdered iron oxide on a properly etched single crystal of iron. The domain boundaries, which are thin regions in which the alignment of the elementary dipoles changes from a certain direction in one domain to an entirely different direction in the other, are the sites of intense but highly localized and nonuniform magnetic fields. The suspended colloidal particles are attracted to these regions. Although the atomic dipoles in the individual domains are completely aligned, the specimen as a whole may have a very small resultant magnetic moment. This is the state of affairs in an unmagnetized iron nail.

As we magnetize a piece of iron by placing it in an external magnetic field, two effects take place. One is a growth in size of the domains that are favorably oriented at the expense of those that are not, as in Fig. 37–15. Second, the direction of orientation of the dipoles within a domain may swing around as a unit, becoming closer to the field direction. Hysteresis comes about because the domain boundaries do not move completely back to their original positions when the external field $\mathbf{B}_0$ is removed.

Two other types of magnetism, closely related to ferromagnetism, are *antiferromagnetism* and *ferrimagnetism* (note spelling). In antiferromagnetic substances, of

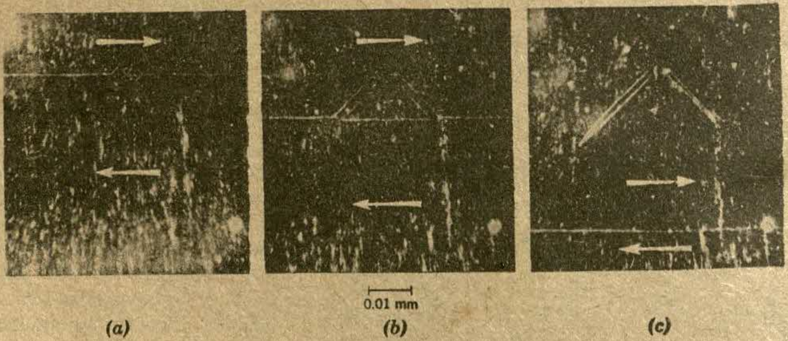


Fig. 37–15 (*a*) A boundary between two domains, with the magnetization in each domain as shown by the white arrows. (*b*) If an external magnetic field pointing from left to right is imposed on the specimen, the upper domain will grow at the expense of the lower. The domain boundary will move down as the elementary dipoles reverse themselves. (*c*) The process continues. The boundary has moved across a region in which there is a crystal imperfection. (Courtesy H. J. Williams, Bell Telephone Laboratories.)

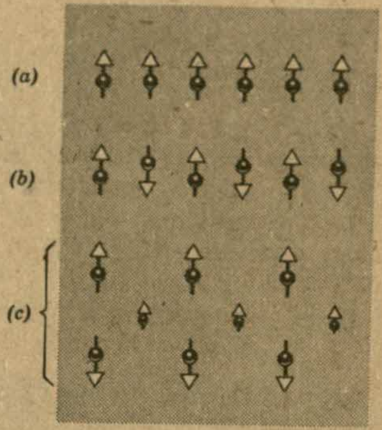


Fig. 37–16 Showing how elementary magnetic dipoles are oriented by the interatomic exchange coupling in (*a*) ferromagnetism, (*b*) antiferromagnetism, and (*c*) ferrimagnetism.

which MnO_2 is an example, the exchange coupling to which we referred on page 834 serves to lock adjacent ions into rigid *antiparallelism* (see Fig. 37–16*b*). Such materials exhibit ~~very little~~ gross external magnetism. However, if they are heated above a certain temperature, called the *Néel temperature*, the exchange coupling ceases to act and the material becomes paramagnetic. In ferrimagnetic substances of which iron ferrite is an example, two different kinds of magnetic ions are present. In iron ferrite the two ions are Fe^{++} and Fe^{+++}. The exchange coupling locks the ions into a pattern like that of Fig. 37–16*c*, in which the external effects are intermediate between ferromagnetism and antiferromagnetism. Here, too, the exchange coupling disappears if the material is heated above a certain characteristic temperature.

37–6 Nuclear Magnetism

Many nuclei have magnetic dipoles, and the possibility arises that a specimen of matter may exhibit gross external magnetic effects associated with its nuclei. However, nuclear magnetic moments are several orders of magnitude smaller than those associated with the electronic motions in an atom or ion. The magnetic moment of an electron associated with its spin, for example, exceeds that of the proton (the nucleus of hydrogen) by a factor of 660.

Gross external effects for nuclear magnetism are smaller than the corresponding (electronic) paramagnetic effects by the *square* of ratios of this order of magnitude, because (*a*), *all else being equal,* the external magnetism is reduced by such a ratio, but (*b*) the very fact that the magnetic dipole moment of the nucleus is smaller means that (see Example 3) the thermal vibrations are proportionally (to a good approximation) more effective in reducing the degree of alignment of the elementary dipoles in an external magnetic field; thus all else is *not* equal and the ratio enters twice.

Techniques such as the Rowland ring (Fig. 37–10) are far too insensitive to detect nuclear magnetism. We describe here a *nuclear resonance technique* by means of which nuclear magnetism can readily reveal itself. This method is also vastly useful for studying paramagnetism, ferromagnetism, antiferromagnetism, and ferrimagnetism, in all of which cases the magnetic effects are associated not with the nuclei but with the atomic electrons. The nuclear-resonance technique was developed in 1946 by E. M. Purcell and his co-workers at Harvard. Simultaneously and independently, F. Bloch and his co-workers at Stanford discovered a very similar method. For these achievements the two physicists received a Nobel prize.

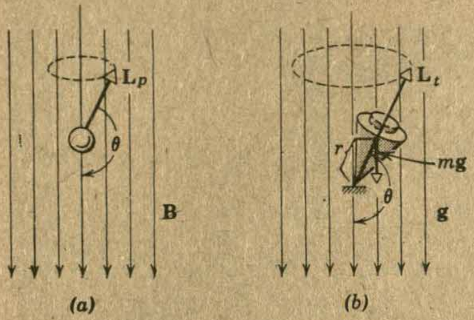


Fig. 37–17 (*a*) A spinning proton precessing in an external magnetic field and (*b*) a spinning top precessing in an external gravitational field. $\mathbf{L}_p$ and $\mathbf{L}_t$ are the two angular momentum vectors.

We focus our attention on the problem of measuring the magnitude μ of the magnetic moment of the proton. In principle, this can be done by placing a specimen containing protons in an external field of magnetic induction **B** and by measuring the energy ($= 2\mu B$) required to turn the protons end for end. A rigorously correct description of the procedures cannot be given without using quantum physics. The description given, although based entirely on classical physics, nevertheless leads to the correct conclusions.

Figure 37–17*a* shows a spinning proton with its axis making an angle θ with a uniform external magnetic field **B**. Figure 37–17*b* shows a spinning top with its axis making an angle θ with a uniform external gravitational field **g**. In each case there is a torque that tends to align the axis of the spinning object with the field. For the proton (Eq. 33–11) it is given by

$$\tau_p = \mu B \sin\theta. \tag{37–17a}$$

For the top it is given by

$$\tau_t = mgr \sin\theta, \tag{37–17b}$$

where r locates the center of mass of the top and m is its mass.

In Section 13–2, we saw that the spinning top processes about a vertical axis with an angular frequency given by

$$\omega_t = \frac{mgr}{L_t}, \tag{37–18a}$$

in which L_t is the spin angular momentum of the top.

The proton, which has a quantized spin angular momentum L_p, will also precess about the direction of the (magnetic) field because of the action of the (magnetic) torque. The student should derive the expression for the frequency of precession, being guided by the derivation of Section 13–2, but using the magnetic torque (Eq. 37–17*a*) instead of the gravitational torque (Eq. 37–17*b*). The relation is

$$\omega_p = \frac{\mu B}{L_p}. \tag{37–18b}$$

▶ **Example 5.** What is the precession frequency of a proton in a magnetic field of 0.5 weber/meter^2?

The quantities μ and L_p in Eq. 37–18*b* are 1.4×10^{-26} amp-m^2 and 0.53×10^{-34} joule-sec. This equation then yields

$$\nu_p = \frac{\omega_p}{2\pi} = \frac{\mu B}{2\pi L_p} = \frac{(1.4 \times 10^{-26}\text{ amp-m}^2)(0.50\text{ weber/meter}^2)}{(2\pi)(0.53 \times 10^{-34}\text{ joule-sec})} = 2.1 \times 10^7\text{ cps.}$$

This frequency (= 21 mc/sec) is in the radio-frequency range. ◀

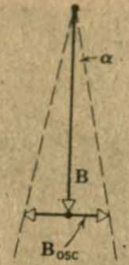


Fig. 37–18 In the nuclear magnetic resonance method a small oscillating magnetic field $\mathbf{B}_{osc}$ is placed at right angles to a steady field **B**.

It is possible to change the energy of any system in periodic motion if we allow an external influence to act on it at the same frequency as that of its motion. This is the familiar *resonance* condition. As an "external influence" for the precessing proton, we use a small alternating magnetic field $\mathbf{B}_{osc}$ arranged to be at right angles to the steady field **B**. This oscillating field combines vectorially with the steady field so that the *resultant* field rocks back and forth between the limits shown by the dashed lines in Fig. 37–18. Typical values for B and for the amplitude of $\mathbf{B}_{osc}$ are 5000 gauss and 1 gauss, respectively, so that the rocking angle α in the figure is quite small. If the angular frequency ω_0 of the oscillating field is chosen equal to the angular precession frequency ω_p of the proton, it turns out that the precessing proton can absorb energy. An increase in energy means an increase in θ in Fig. 37–17*a*.

The resonance condition

$$\omega_0 = \frac{\mu B}{L_p} \tag{37-19}$$

can be used to measure μ. We place the spinning proton in a known field **B**, apply a "perturbing field" at right angles to it, and vary the angular frequency ω_0 of this perturbing field until resonance occurs. It is possible to tell when Eq. 37–19 is satisfied because, at resonance, many spinning protons will tend to turn end for end in the field, absorbing energy which can be detected by appropriate electronic techniques.

Figure 37–19 is a schematic diagram of an experimental arrangement. The protons, present as hydrogen nuclei in a small vial V of water, are immersed in a strong steady magnetic field caused by the electromagnet whose pole faces N and S are shown. A rapidly alternating current in the small coil C provides the (horizontal) weak, perturbing magnetic field $\mathbf{B}_{osc}$. This current is provided by a radio-frequency oscillato

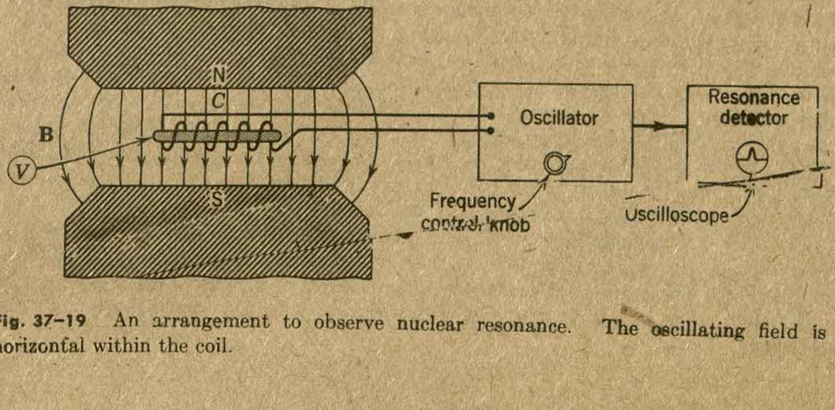


Fig. 37–19 An arrangement to observe nuclear resonance. The oscillating field is horizontal within the coil.

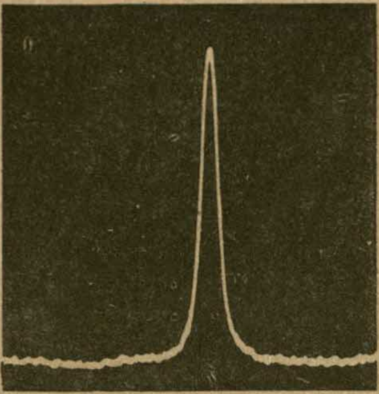


Fig. 37-20 An oscilloscope photograph of a proton resonance peak showing energy absorbed from the oscillator versus oscillator frequency. (From Bloembergen et al., *Phys. Rev.*, 73, 679.)

whose angular frequency ω_0 can be varied; an electronic "resonance detector," also connected to the oscillator, serves to indicate when energy is being drained from the oscillator and used to "flip the protons." In principle, the oscillator angular frequency ω_0 is varied until the resonance detector shows that Eq. 37-19 is satisfied (see Fig. 37-20). The magnetic moment μ can then be determined by measuring B and ω_0. Surprisingly enough, magnetic moments can be measured in this and similar ways to a much greater accuracy than we can measure μ for a bar magnet. For the proton we have

$$\mu_p = 1.41044 \times 10^{-26} \text{ amp-m}^2$$

37-7 Three Magnetic Vectors

In Chapter 30 we saw that if a dielectric is placed in an electric field polarization charges will appear on its surface. These surface charges, which find their origin in the elementary electric dipoles (permanent or induced) that make up the dielectric, set up a field of their own that modifies the original field. For the simple case discussed in Chapter 30—a dielectric slab in a parallel-plate capacitor—this complication can readily be handled in terms of the electric field strength vector **E** and some knowledge of the electric properties of the slab material, such as the dielectric constant. For more complex problems we asserted that it was useful to introduce two other (subsidiary) electric vectors, the *electric polarization* **P** and the *electric displacement* **D**. Table 30-2 shows some of the characteristics of these three vectors.

In magnetism we find a similar situation. If magnetic materials are placed in a field of induction, the elementary magnetic dipoles (permanent or induced) will act to set up a field of induction of their own that will modify the original field. For the simple case discussed in this chapter—a Rowland ring with a ferromagnetic core—this complication can readily be handled in terms of the magnetic induction vector **B** and some knowledge of the magnetic properties of the ring material, such as that provided by Fig. 37-9. For more complex problems we find it useful to introduce two other (subsidiary) magnetic vectors, the *magnetization* **M** and the *magnetic field strength* **H**. We do so largely so that the student who takes a second course in electromagnetism will have some familiarity with them.

Consider a Rowland ring carrying a current i_0 in its windings and designed so that its core, assumed to be iron, can be removed. The magnetic induction B, measured by the methods of Section 37-5, will be much greater when the core is in place than when it is not, assuming that the current in the windings remains unchanged.

Physically we can understand the large value of B in the iron core in terms of the alignment of the elementary dipoles in the iron. A hypothetical slice out of the iron core, as in Fig. 37-21*b*, has a magnetic moment $d\boldsymbol{\mu}$ equal to the vector sum of all of the elementary dipoles contained in it. We define our first subsidiary vector, the *magnetization* **M**, as the magnetic moment per unit volume of the core material. For the slice of Fig. 37-21*b* we have

$$d\boldsymbol{\mu} = \mathbf{M}(A\,dl),$$

where $(A\,dl)$ is the volume of the slice, A being the cross section of the core.

When we discussed Ampère's law in Chapter 34, we assumed that no magnetic materials were present. If we apply this law, namely

$$\oint \mathbf{B}\cdot d\mathbf{l} = \mu_0 i, \tag{37-20}$$

to the circular path of integration shown in Fig. 37-21*a*, we obtain

$$(B)(2\pi r_0) = \mu_0(N_0 i_0), \tag{37-21}$$

in which r_0 is the mean radius of the core, N_0 is the number of turns, and i_0 is the current in each turn. We see at once that Ampère's law, in the form of Eq. 37-20,

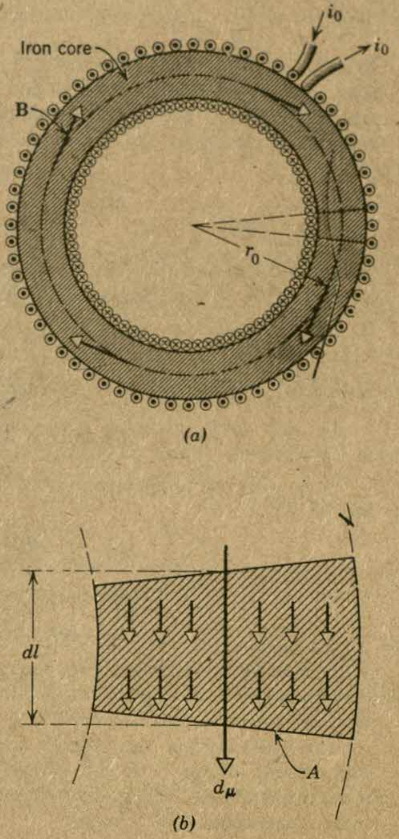


Fig. 37-21 (*a*) A Rowland ring with an iron core. (*b*) A slice of the core, showing its magnetic moment $d\boldsymbol{\mu}$ caused by the alignment of the elementary magnetic dipoles in the iron.

is not valid when magnetic materials are present. Equation 37–21 predicts that, since the right side is the same whether or not the core is in place, the induction B should also be the same, a prediction not in accord with experiment.

We can increase B in the absence of the iron core to the value that it has when the core is in place if we increase the current in the windings by an amount $i_{M,0}$. The magnetization of the iron core is thus *equivalent in its effect on* **B** to such a hypothetical current increase. We choose to modify Ampère's law by arbitrarily inserting a *magnetizing current* term i_M on the right, obtaining

$$\oint \mathbf{B}\cdot d\mathbf{l} = \mu_0(i + i_M). \tag{37-22}$$

If we give i_M a suitable value when the iron core is in place, it is clear that Ampère's law, in this new form, can remain valid. It remains to relate this (largely hypothetical) magnetizing current to something more physical, the magnetization **M**.*

Applying Eq. 37–22 to the iron ring of Fig. 37–21a yields

$$(B)(2\pi r_0) = \mu_0(N_0 i_0) + \mu_0(N_0 i_{M,0}). \tag{37-23}$$

We can relate $i_{M,0}$ to the magnetization **M** if we recall (Eq. 33–10) that the magnetic moment of a magnetic dipole in the form of a current loop is given by

$$\mu = NiA,$$

where N is the number of turns in the loop, i is the loop current, and A is the loop area. Let us use this equation to find what increase $i_{M,0}$ in current in the windings around the slice of Fig. 37–21b would produce a magnetic moment equivalent to that actually produced by the alignment of elementary dipoles in the slice. We have

$$M(A\,dl) = \left(N_0\frac{dl}{2\pi r_0}\right)(i_{M,0})(A),$$

the quantity in the first parentheses on the right being the number of turns associated with the slice of thickness dl. This reduces to

$$N_0 i_{M,0} = M(2\pi r_0). \tag{37-24}$$

Substituting this into Eq. 37–23 yields

$$(B)(2\pi r_0) = \mu_0(N_0 i_0) + \mu_0(M)(2\pi r_0). \tag{37-25}$$

We now choose to generalize from the special case of the Rowland ring by writing Eq. 37–25 as

$$\oint \mathbf{B}\cdot d\mathbf{l} = \mu_0 i + \mu_0 \oint \mathbf{M}\cdot d\mathbf{l}$$

or

$$\oint \left(\frac{\mathbf{B} - \mu_0\mathbf{M}}{\mu_0}\right)\cdot d\mathbf{l} = i.$$

---

* It is possible to give reality to the magnetizing current by viewing it as a real current that flows around the magnet at its surface, being the resultant macroscopic effect of all the microscopic current loops that constitute the atomic electron orbits. This *Amperian current* viewpoint however does not take the magnetization due to electron spin readily into account. Since we do not attempt to measure magnetizing currents experimentally, other than through their (postulated) magnetic effects, we prefer to view the magnetizing current as a convenient formalism.

The quantity $(\mathbf{B} - \mu_0\mathbf{M})/\mu_0$ occurs so often in magnetic situations that we give it a special name, the *magnetic field strength* **H**, or

$$\mathbf{H} = \frac{\mathbf{B} - \mu_0\mathbf{M}}{\mu_0}$$

which we write as

$$\mathbf{B} = \mu_0\mathbf{H} + \mu_0\mathbf{M}. \qquad (37\text{–}26)$$

Ampère's law can now be written in the simple form

$$\oint \mathbf{H}\cdot d\mathbf{l} = i, \qquad (37\text{–}27)$$

which holds in the presence of magnetic materials and in which i is the *true current only*, that is, it does not include the magnetizing current. This reminds us that the electric displacement vector **D** permitted us to write Gauss's law for the case in which dielectric materials are present, in a form involving free charges only, that is, not polarization charges; see Table 30–2.

We state without proof (see Problems 12 and 13) that at a boundary between two media (1) the component of **H** tangential to the surface has the same value on each side of the surface * and (2) the component of **B** perpendicular to the surface has the same value on each side of the surface. These *boundary conditions* are of great value in solving complex problems.

To find H in our Rowland ring, let us apply Ampère's law in the generalized form of Eq. 37–27. We have

$$(H)(2\pi r_0) = N_0 i_0,$$

where i_0 is the (true) current in the windings. This gives

$$H = \left(\frac{N_0}{2\pi r_0}\right) i_0 = n i_0, \qquad (37\text{–}28)$$

in which n is the number of turns per unit length. Since we have not introduced any information describing the core into Eq. 37–27, the value of H computed from Eq. 37–28 is independent of the core material.

B can be measured experimentally by the method of Section 37–5 and M can then be calculated from Eq. 37–26. The student should note in passing (see Eq. 37–15) that the abscissa B_0 in Fig. 37–9 is proportional to H $(= \mu_0 H)$, the ordinate being proportional to B. Curves such as this and that of Fig. 37–12 are called *B-H curves.*

Let us assume that we have made measurements of **H**, **B**, and **M** for a wide variety of magnetic materials, using either the technique described or an equivalent one. For *paramagnetic* and *diamagnetic* materials we would find, as an experimental result, that **B** is directly proportional to **H**, or

$$\mathbf{B} = \kappa_m\mu_0\mathbf{H}, \qquad (37\text{–}29)$$

in which κ_m, the *permeability* of the magnetic medium, is a constant for a given temperature and density of the material. Eliminating **B** between Eqs. 37–29 and 37–26 allows us to write

$$\mathbf{M} = (\kappa_m - 1)\mathbf{H}, \qquad (37\text{–}30)$$

which is another expression of the linear or proportional character of paramagnetic and diamagnetic materials.

---

* Assuming that there are no true currents at the surface, as there are in the Rowland ring of Fig. 37–21a, for example.

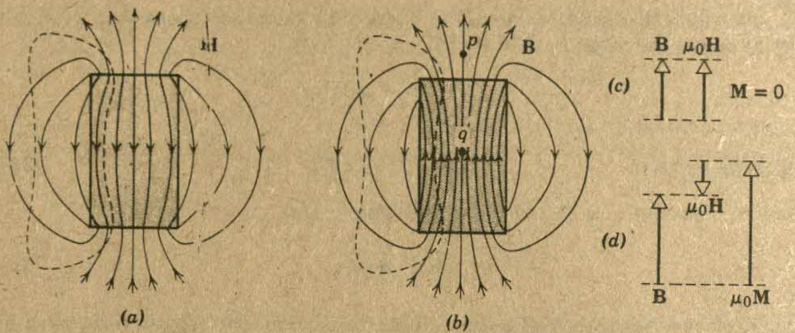


Fig. 37–22 (*a*) The lines of **H** and (*b*) the lines of **B** for a permanent magnet. Note that the lines of **H** change direction at the boundary. The closed dashed curves are paths of integration around which Ampère's law may be applied. The relation $\mathbf{B} = \mu_0\mathbf{H} + \mu_0\mathbf{M}$ is shown to be satisfied for (*c*) a particular outside point *p* and (*d*) a particular inside point *q*.

For a vacuum, in which there are no magnetic dipoles present to be aligned, the magnetization **M** must be zero. Putting $\mathbf{M} = 0$ in Eq. 37–26 leads to

$$\mathbf{B} = \mu_0\mathbf{H} \qquad \text{(a vacuum).} \qquad (37\text{–}31)$$

Comparison with Eq. 37–29 shows that a vacuum must be described by $\kappa_m = 1$. Equation 37–30 verifies that the magnetization vanishes if we put κ_m equal to unity. For paramagnetic materials κ_m is slightly greater than unity. For diamagnetic materials it is slightly less than unity; Eq. 37–30 shows that this requires **M** and **H** to be oppositely directed, a fact discussed at length in Section 37–4.

In ferromagnetic materials the relationship between **B** and **H** is far from linear, as Figs. 37–9 and 37–12 show. Experimentally, κ_m proves to be a function not only of the value of H but also, because of hysteresis, of the magnetic and thermal history of the specimen.*

An interesting special case of ferromagnetism is the permanent magnet, for which **H**, **M**, and **B** all have nonvanishing values inside the magnet even though there is no true current. Figure 37–22 shows typical lines of **B** and **H** associated with such a magnet. The lines of **B** may be drawn as continuous loops, the boundary condition (2), mentioned above, being satisfied where the lines enter and leave the magnet. Equation 37–22 shows that the vector **B** is associated with the *total* current, both true and magnetizing. In Fig. 37–22*a* $\oint \mathbf{B}\cdot d\mathbf{l}$ around any loop such as that shown by the dashed curve is not zero and must be associated with a hypothetical magnetizing current i_M imagined to circulate around the magnet at its surface; actual or true currents (i) do not exist in this problem. Figure 37–22*b* shows that **H** reverses direction at the boundary. Since **H** (see Eq. 37–27) is associated with true currents only, we must have $\oint \mathbf{H}\cdot d\mathbf{l}$ around any loop such as that shown by the dashed lines.

---

* In the dielectric case there are waxy materials, called *ferroelectrics*, for which the relationship between **D** and **E** is nonlinear, which exhibit hysteresis, and from which quasi-permanent electric dipoles (*electrets*) can be constructed. However, most commonly useful dielectric materials are linear whereas most commonly useful magnetic materials are nonlinear.

The reversal of **H** at the boundary makes this possible. Note that **M** and **H** point in opposite directions within the magnet. Table 37-1 summarizes the properties of the three vectors **B**, **H**, and **M**.

Table 37-1

THREE MAGNETIC VECTORS

Name	Symbol	Associated with	Boundary Condition
Magnetic induction	**B**	All currents	Normal component continuous
Magnetic field strength	**H**	True currents only	Tangential component continuous †
Magnetization (magnetic dipole moment per unit volume)	**M**	Magnetization currents only	Vanishes in a vacuum

Defining equations for **B**	$\mathbf{F} = q\mathbf{v} \times \mathbf{B}$ or $= i\mathbf{l} \times \mathbf{B}$	Eq. 33-3*a* Eq. 33-6*a*
General relation among the three vectors	$\mathbf{B} = \mu_0\mathbf{H} + \mu_0\mathbf{M}$	Eq. 37-26
Ampère's law when magnetic materials are present	$\oint \mathbf{H}\cdot d\mathbf{l} = i$ (i = true current only)	Eq. 37-27
Empirical relations for certain magnetic materials *	$\mathbf{B} = \kappa_m\mu_0\mathbf{H}$ $\mathbf{M} = (\kappa_m - 1)\mathbf{H}$	Eq. 37-29 Eq. 37-30

* For paramagnetic and diamagnetic materials only, if κ_m is to be independent of **H**.

† Assuming no true currents exist at the boundary.

▶ **Example 6.** *In the Rowland ring the (true) current i_0 in the windings is 2.0 amp and the number of turns per unit length (n) in the toroid is 10 turns/cm. B, measured by the technique of Section 37-5, is 1.0 weber/meter^2. Calculate (a) H, (b) M, and (c) the magnetizing current $i_{M,0}$ both when the core is in place and when it is removed. (d) For these particular operating conditions, what is κ_m?*

(*a*) H is independent of the core material and may be found from Eq. 37-28:

$$\begin{aligned} H &= ni \\ &= (10 \text{ turns/cm})(2.0 \text{ amp}) \\ &= 2.0 \times 10^3 \text{ amp/meter}. \end{aligned}$$

(b) M is zero when the core is removed. With the core in place, we may solve Eq. 37–26 for **M**, obtaining for the magnitude of **M**

$$\begin{aligned} M &= \frac{B - \mu_0 H}{\mu_0} \\ &= \frac{(1.0\ \text{weber/meter}^2) - (4\pi \times 10^{-7}\ \text{weber/amp-m})(2.0 \times 10^3\ \text{amp/meter})}{(4\pi \times 10^{-7}\ \text{weber/amp-m})} \\ &= 7.9 \times 10^5\ \text{amp/meter}. \end{aligned}$$

(c) The effective magnetizing current follows from Eq. 37–24:

$$\begin{aligned} i_{M,0} &= M\left(\frac{2\pi r_0}{N_0}\right) = \frac{M}{n} \\ &= \frac{7.9 \times 10^5\ \text{amp/meter}}{2.0 \times 10^3\ \text{turns/meter}} \\ &= 390\ \text{amp}. \end{aligned}$$

An *additional* current of this amount in the windings would produce the same value of B in the absence of a core as that obtained, by alignment of the elementary dipoles, with the core in place.

(d) The permeability can be found from Eq. 37–29, or

$$\begin{aligned} \kappa_m &= \frac{B}{\mu_0 H} \\ &= \frac{1.0\ \text{weber/meter}^2}{(4\pi \times 10^{-7}\ \text{weber/amp-m})(2.0 \times 10^3\ \text{amp/meter})} \\ &= 397. \end{aligned}$$

We emphasize that this value of κ_m holds only for the special conditions of this experiment. ◀

QUESTIONS

1. Two iron bars are identical in appearance. One is a magnet and one is not. How can you tell them apart? You are not permitted to suspend either bar as a compass needle or to use any other apparatus.

2. How could you reverse the magnetism of a compass needle?

3. Two iron bars always attract, no matter the combination in which their ends are brought near each other. Can you conclude that one of the bars must be unmagnetized?

4. If we sprinkle iron filings on a particular bar magnet, they cling both to the ends *and to the middle*. Sketch roughly the lines of **B**, both outside and inside the magnet.

5. The earth is a huge magnetic dipole. (a) Is the magnetic pole in the Northern Hemisphere a north or a south magnetic pole? (b) In the Northern Hemisphere do the magnetic lines of force associated with the earth's magnetic field point toward the earth's surface or away from it?

6. Cosmic rays are charged particles that strike our atmosphere from some external source. We find that more low-energy cosmic rays reach the earth at the north and south magnetic poles than at the (magnetic) equator. Why is this so?

7. How might the magnetic dipole moment of the earth be measured?

8. Give three reasons for believing that the flux Φ_B of the earth's magnetic field is greater through the boundaries of Alaska than through those of Texas.

9. The neutron, which has no charge, has a magnetic dipole moment. Is this possible on the basis of classical electromagnetism, or does this evidence alone indicate that classical electromagnetism has broken down?

10. Is the magnetization at saturation for a paramagnetic substance very much different from that for a saturated ferromagnetic substance of about the same size?

11. Explain why a magnet attracts an unmagnetized iron object such as a nail.

12. Does any net force or torque act on (*a*) an unmagnetized iron bar or (*b*) a permanent bar magnet when placed in a uniform magnetic field?

13. A nail is placed at rest on a smooth table top near a strong magnet. It is released and attracted to the magnet. What is the source of the kinetic energy it has just before it strikes the magnet?

14. The magnetization induced in a given diamagnetic sphere by a given external magnetic field does not vary with temperature, in sharp contrast to the situation in paramagnetism. Is this understandable in terms of the description that we have given of the origin of diamagnetism?

15. Compare the magnetization curves for a paramagnetic substance (Fig. 37–6) and for a ferromagnetic substance (Fig. 37–9). What would a similar curve for a diamagnetic substance look like? Do you think that it would show saturation effects in strong applied fields (say 10 weber/meter^2)?

16. Distinguish between the precession frequency and the cyclotron frequency of a proton in a magnetic field.

17. In our discussion of nuclear magnetism we said that energy absorption occurs because the dipoles are turned end for end. However, a given dipole might initially be lined up either with the field or against it. In the first case there would be an *absorption* of energy, but in the second case there would be a *release* of energy, each amount being $2\mu B$. Why do we observe a *net* absorption? These two events would seem to cancel.

18. Discuss similarities and differences in Tables 30–2 and 37–1.

PROBLEMS

1. The earth has a magnetic dipole moment of 6.4×10^{21} amp-m^2. (*a*) What current would have to be set up in a single turn of wire going around the earth at its magnetic equator if we wished to set up such a dipole? (*b*) Could such an arrangement be used to cancel out the earth's magnetism at points in space well above the earth's surface? (*c*) On the earth's surface?

2. Assume that the electron is a small sphere of radius R, its charge and mass being spread uniformly throughout its volume. Such an electron has a "spin" angular momentum L of 0.53×10^{-34} joule-sec and a magnetic moment μ of 9.3×10^{-24} amp-m^2. Show that $e/m = 2\mu/L$. Is this prediction in agreement with experiment? (Hint: The spherical electron must be divided into infinitesimal current loops and an expression for the magnetic moment found by integration. This model of the electron is too mechanistic to be in the spirit of the modern quantum view of this particle.)

3. Calculate (*a*) the electric field strength and (*b*) the magnetic induction at a point 1.0 A (one angstrom unit) away from a proton, measured along its axis of spin. The magnetic moment of the proton is 1.4×10^{-26} amp-m^2.

4. A Rowland ring is formed of ferromagnetic material. It is circular in cross section, with an inner radius of 5.0 cm and an outer radius of 6.0 cm and is wound with 400 turns of wire. (*a*) What current must be set up in the windings to attain $B_0 = 2 \times 10^{-4}$ weber/meter^2 in Fig. 37–9? (*b*) A secondary coil wound around the toroid has 50 turns and has a resistance of 8.0 ohms. If, for this value of B_0, we have $B_M = 800B_0$, how much charge moves through the secondary coil when the current in the toroid windings is turned on?

5. The dipole moment associated with an atom of iron in an iron bar is 1.8×10^{-23} amp-m^2. Assume that all the atoms in the bar, which is 5 cm long and has a cross-sec-

tional area of 1 cm^2, have their dipole moments aligned. (*a*) What is the dipole moment of the bar? (*b*) What torque must be exerted to hold this magnet at right angles to an external field of 15,000 gauss?

6. Can you give an explanation of diamagnetism based on Faraday's law of induction? In Figs. 37–7*a* and *b*, for example, what inductive effects can be expected as the magnetic field is built up from zero to the final value **B**?

7. Prove that $\Delta\omega \ll \omega_0$ in Eq. 37–13.

8. Show that, classically, a spinning positive charge will have a spin magnetic moment that points in the same direction as its spin angular momentum.

9. Assume that the nuclei (protons) in 1 gm of water *could all be aligned*. What magnetic induction B would be produced 5.0 cm from the sample, along its alignment axis?

10. It is possible to measure e/m for the electron by measuring (*a*) the cyclotron frequency ν_c of electrons in a given magnetic field and (*b*) the precession frequency ν_p of protons in the same field. Show that the relation is

$$\frac{e}{m} = \frac{\nu_c}{\nu_p}\frac{\mu_s}{L_s}.$$

Since μ_s and L_s for the proton are accurately known, this experiment gives us our most precise value of e/m today.

11. *Dipole-dipole interaction.* The exchange coupling mentioned in Section 37–5 as being responsible for ferromagnetism is *not* the mutual magnetic interaction energy between two elementary magnetic dipoles. To show this (*a*) compute B a distance a ($= 1.0$ A) away from a dipole of moment μ ($= 1.8 \times 10^{-23}$ amp-m^2); (*b*) compute the energy ($= 2\mu B$) required to turn a second similar dipole end for end in this field. What do you conclude about the strength of this dipole-dipole interaction? Compare with the results of Example 3. (Note: for the same distance, the field in the median plane of a dipole is only half as large as on the axis; see Eq. 37–2.)

12. *Boundary condition for* **B**. Prove that at the boundary between two media the normal component of **B** has the same value on each side of the surface. (Hint: Construct a closed Gaussian surface shaped like a flat pillbox with one face in each medium and apply Gauss's law for magnetism.)

13. *Boundary condition for* **H**. Prove that at the boundary between two media the tangential component of **H** has the same value on each side of the surface, assuming that here is no current at the surface. (Hint: Construct a closed rectangular loop, the two posite longer sides being parallel to the surface, with one side in each medium. Use père's law in the form that applies when magnetic materials are present.)

14. The magnetic energy density can be shown to be given in its most general form as

$$u_B = \tfrac{1}{2}\mathbf{B}\cdot\mathbf{H}.$$

Does this reduce to a familiar result for a vacuum?

Electromagnetic Oscillations

CHAPTER 38

38–1 *LC* Oscillations

The *LC* system of Fig. 38–1 resembles a mass-spring system (see Fig. 8–4) in that, among other things, each system has a characteristic frequency of oscillation. To see this, we assume that initially the capacitor C in Fig. 38–1a carries a charge q_m and the current i in the inductor is zero. At this instant the energy stored in the capacitor is given by Eq. 30–25, or

$$U_E = \frac{1}{2}\frac{{q_m}^2}{C}. \tag{38–1}$$

The energy stored in the inductor, given by

$$U_B = \tfrac{1}{2}Li^2, \tag{38–2}$$

is zero because the current is zero. The capacitor now starts to discharge through the inductor, positive charge carriers moving counterclockwise, as shown in Fig. 38–1b. This means that a current i, given by dq/dt and pointing down in the inductor, is established.

As q decreases the energy stored in the electric field in the capacitor also decreases. This energy is transferred to the magnetic field that appears around the inductor because of the current i that is building up there. Thus the electric field decreases, the magnetic field builds up, and energy is transferred from the former to the latter.

At a time corresponding to Fig. 38–1c, all the charge on the capacitor will have disappeared. The electric field in the capacitor will be zero, the energy stored there having been transferred entirely to the magnetic field of the

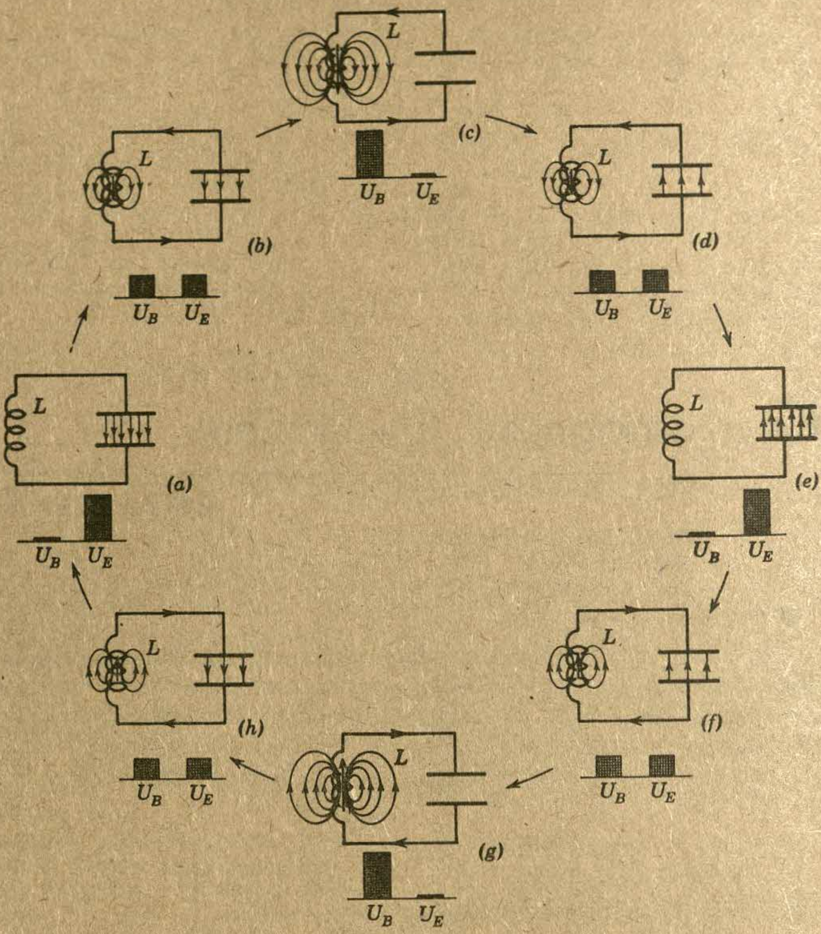


Fig. 38–1 Showing eight stages in a cycle of oscillation of an LC circuit. The bar graphs below each figure show the stored magnetic and electric potential energy. The vertical arrows on the inductor axis show the current. The student should compare this figure in detail with Fig. 8–4, to which it exactly corresponds.

inductor. According to Eq. 38–2, there must then be a current—and indeed one of maximum value—in the inductor. Note that even though q equals zero the current (which is dq/dt) is *not* zero at this time.

The large current in the inductor in Fig. 38–1*c* continues to transport positive charge from the top plate of the capacitor to the bottom plate, as shown in Fig. 38–1*d*; energy now flows from the inductor back to the capacitor as the electric field builds up again. Eventually, the energy will have been transferred completely back to the capacitor, as in Fig. 38–1*e*. The situation of Fig. 38–1*e* is like the initial situation, except that the capacitor is charged in the opposite direction.

The capacitor will start to discharge again, the current now being clockwise, as in Fig. 38-1*f*. Reasoning as before, we see that the circuit eventually returns to its initial situation and that the process continues at a definite frequency ν (measured, say, in cycles/sec) to which corresponds a definite *angular* frequency ω ($= 2\pi\nu$ and measured, say, in radians/sec). Once started, such *LC* oscillations (in the ideal case described, in which the circuit contains no resistance) continue indefinitely, energy being shuttled back and forth between the electric field in the capacitor and the magnetic field in the inductor. Any configuration in Fig. 38-1 can be set up as an initial condition. The oscillations will then continue from that point, proceeding clockwise around the figure. The student should compare these oscillations carefully with those of the mass-spring system described in Fig. 8-4.

To measure the charge q as a function of time, we can measure the variable potential difference $V_C(t)$ that exists across capacitor C. The relation

$$V_C = \left(\frac{1}{C}\right)q$$

shows that V_C is proportional to q. To measure the current we can insert a small resistance R in the circuit and measure the potential difference across it. This is proportional to i through the relation

$$V_R = (R)i.$$

We assume here that R is so small that its effect on the behavior of the circuit is negligible. Both q and i, or more correctly V_C and V_R, which are proportional to them, can be displayed on a cathode-ray oscilloscope. This instrument can plot automatically on its screen graphs proportional to $q(t)$ and $i(t)$, as in Fig. 38-2.

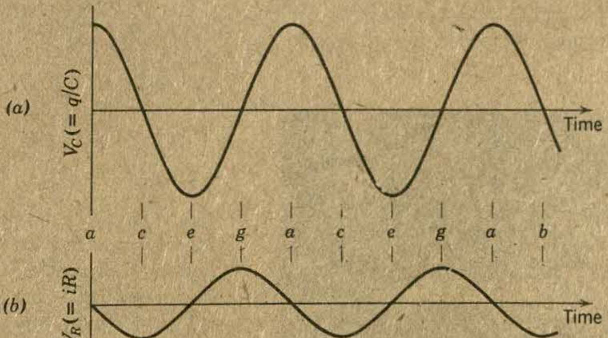


Fig. 38-2 A drawing of an oscilloscope screen showing potential differences proportional to (*a*) the charge and (*b*) the current, in the circuit of Fig. 38-1, as a function of time. The letters indicate corresponding phases of oscillation in that figure. Note that because $i = dq/dt$ the lower curve is proportional to the derivative of the upper.

▶ **Example 1.** A 1.0-μf capacitor is charged to 50 volts. The charging battery is then disconnected and a 10-mh coil is connected across the capacitor, so that LC oscillations occur. What is the maximum current in the coil? Assume that the circuit contains no resistance.

The maximum stored energy in the capacitor must equal the maximum stored energy in the inductor, from the conservation-of-energy principle. This leads, from Eqs. 38–1 and 38–2, to

$$\frac{1}{2}\frac{q_m^2}{C} = \tfrac{1}{2}Li_m^2,$$

where i_m is the *maximum* current and q_m is the *maximum* charge. Note that the maximum current and the maximum charge do not occur at the same time but one-fourth of a cycle apart; see Figs. 38–1 and 38–2. Solving for i_m and substituting CV_0 for q_m gives

$$i_m = V_0\sqrt{\frac{C}{L}} = (50 \text{ volts})\sqrt{\frac{1.0 \times 10^{-6} \text{ farad}}{10 \times 10^{-3} \text{ henry}}} = 0.50 \text{ amp.}$$ ◀

In an actual LC circuit the oscillations will not continue indefinitely because there is always some resistance present that will drain away energy by Joule heating. The oscillations, once started, will die away, as in Fig. 38–3. This figure should be compared to Fig. 15–19, which shows the decay of the mechanical oscillations of a mass-spring system caused by frictional damping.

It is possible to have sustained electromagnetic oscillations if arrangements are made to supply, automatically and periodically (once a cycle, say), enough energy from an outside source to compensate for that lost to Joule heat. We are reminded of a clock escapement, which is a device for feeding energy from a spring or a falling weight into an oscillating pendulum, thus compensating for frictional losses that would otherwise cause the oscillations to die away. Oscillators whose frequency ν may be varied between certain limits are commercially available as packaged units over a wide range of frequencies, extending from low audio-frequencies (lower than 10 cycles/sec) to microwave frequencies (higher than 10^{10} cycles/sec).

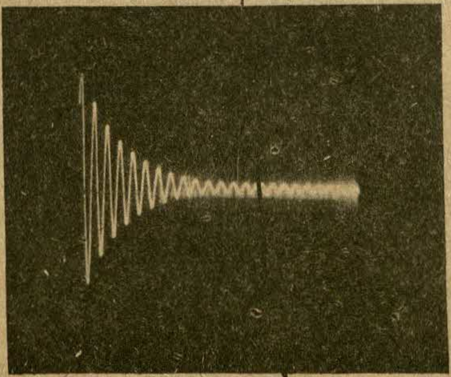


Fig. 38–3 A photograph of an oscilloscope trace showing how the oscillations in an LRC circuit die away because energy is lost to Joule heat in the resistor. The figure is a plot of the potential difference across the resistor as a function of time.

38–2 Analogy to Simple Harmonic Motion

Figure 8–4 shows that in a mass-spring system performing simple harmonic motion, as in an oscillating *LC* circuit, two kinds of energy occur. One is potential energy of the compressed or extended spring; the other is kinetic energy of the moving mass. These are given by the familiar formulas in the first column of Table 38–1. The table suggests that a capacitor is in some formal way like a spring and an inductor is like a mass and that certain electromagnetic quantities "correspond" to certain mechanical ones, namely,

Table 38–1

SOME ENERGY FORMULAS

Mechanical		Electromagnetic	
spring	$U_P = \frac{1}{2}kx^2$	capacitor	$U_E = \frac{1}{2}\frac{q^2}{C}$
mass	$U_K = \frac{1}{2}mv^2$	inductor	$U_B = \frac{1}{2}Li^2$
$v = dx/dt$		$i = dq/dt$	

q corresponds to x,
i corresponds to v,
C corresponds to $1/k$,
L corresponds to m.

Comparison of Fig. 38–1, which shows the oscillations of the *LC* circuit, with Fig. 8–4, which shows the oscillations in a mass-spring system, indicates how close the correspondence is. Note how v and i correspond in the two figures; also x and q. Note, too, how in each case the energy alternates between two forms, magnetic and electric for the *LC* system, and kinetic and potential for the mass-spring system.

In Section 15–3 we saw that the natural angular frequency of oscillation of an undamped mass-spring system is

$$\omega = 2\pi\nu = \sqrt{\frac{k}{m}}.$$

The method of correspondences suggests that to find the natural frequency for the *LC* circuit k should be replaced by $1/C$ and m by L, obtaining

$$\omega = 2\pi\nu = \sqrt{\frac{1}{LC}}. \qquad (38\text{–}3)$$

This formula is indeed correct, as we show in the next section.

38–3 Electromagnetic Oscillations—Quantitative

We now derive an expression for the frequency of oscillation of an LC circuit, our derivation being based on the conservation-of-energy principle. The total energy U present at any instant in an oscillating LC circuit is given by

$$U = U_B + U_E = \tfrac{1}{2}Li^2 + \frac{1}{2}\frac{q^2}{C},$$

which expresses the fact that at any arbitrary time the energy is stored partly in the magnetic field in the inductor and partly in the electric field in the capacitor. If we assume the circuit resistance to be zero, there is no energy transfer to Joule heat and U remains constant with time, even though i and q vary. In more formal language, dU/dt must be zero. This leads to

$$\frac{dU}{dt} = \frac{d}{dt}\left(\tfrac{1}{2}Li^2 + \frac{1}{2}\frac{q^2}{C}\right) = Li\frac{di}{dt} + \frac{q}{C}\frac{dq}{dt} = 0. \tag{38–4}$$

Now, q and i are not independent variables, being related by

$$i = \frac{dq}{dt}.$$

Differentiating yields

$$\frac{di}{dt} = \frac{d^2q}{dt^2}$$

Substituting these two expressions into Eq. 38–4 leads to

$$L\frac{d^2q}{dt^2} + \frac{1}{C}q = 0. \tag{38–5}$$

This is the differential equation that describes the oscillations of a (resistanceless) LC circuit. To solve it, the student should note that Eq. 38–5 is mathematically of exactly the same form as Eq. 15–6,

$$m\frac{d^2x}{dt^2} + kx = 0, \tag{15–3}$$

which is the differential equation for the mass-spring oscillations. Fundamentally, it is by comparing these two equations that the correspondences on p. 947 arise.

The solution of Eq. 15–6 proved to be

$$x = A\cos(\omega t + \phi), \tag{15–5}$$

where A $(= x_m)$ is the amplitude of the motion and ϕ is an arbitrary *phase angle*. Since q corresponds to x, we can write the solution of Eq. 38–5 as

$$q = q_m\cos(\omega t + \phi), \tag{38–6}$$

where ω is the still unknown angular frequency of the electromagnetic oscillations.

We can test whether Eq. 38-6 is indeed a solution of Eq. 38-5 by substituting it and its second derivative in that equation. To find the second derivative, we write

$$\frac{dq}{dt} = i = -\omega q_m \sin(\omega t + \phi) \tag{38-7}$$

and

$$\frac{d^2q}{dt^2} = -\omega^2 q_m \cos(\omega t + \phi).$$

Substituting q and d^2q/dt^2 into Eq. 38-5 yields

$$-L\omega^2 q_m \cos(\omega t + \phi) + \frac{1}{C} q_m \cos(\omega t + \phi) = 0.$$

Canceling $q_m \cos(\omega t + \phi)$ and rearranging leads to

$$\omega = \sqrt{\frac{1}{LC}}.$$

Thus, if ω is given the constant value $1/\sqrt{LC}$, Eq. 38-6 is indeed a solution of Eq. 38-5. This expression for ω agrees with Eq. 38-3, which was arrived at by the method of correspondences.

The phase angle ϕ in Eq. 38-6 is determined by the conditions that prevail at $t = 0$. If the initial condition is as represented by Fig. 38-1*a*, then we put $\phi = 0$ in order that Eq. 38-6 may predict $q = q_m$ at $t = 0$. What initial condition in Fig. 38-1 is implied if we select $\phi = 90°$?

▶ **Example 2.** (*a*) In an oscillating *LC* circuit, what value of charge, expressed in terms of the maximum charge, is present on the capacitor when the energy is shared equally between the electric and the magnetic field? (*b*) How much time is required for this condition to arise, assuming the capacitor to be fully charged initially? Assume that $L = 10$ mh and $C = 1.0\ \mu$f.

(*a*) The stored energy and the *maximum* stored energy in the capacitor are, respectively,

$$U_E = \frac{q^2}{2C} \quad \text{and} \quad U_{E,m} = \frac{{q_m}^2}{2C}.$$

Substituting $U_E = \frac{1}{2}U_{E,m}$ yields

$$\frac{q^2}{2C} = \frac{1}{2}\frac{{q_m}^2}{2C} \quad \text{or} \quad q = \frac{1}{\sqrt{2}} q_m.$$

(*b*) To find the time, we write, assuming $\phi = 0$ in Eq. 38-6,

$$q = q_m \cos \omega t = \frac{1}{\sqrt{2}} q_m,$$

which leads to

$$\omega t = \cos^{-1} \frac{1}{\sqrt{2}} = \frac{\pi}{4} \quad \text{or} \quad t = \frac{\pi}{4\omega}.$$

The angular frequency ω is found from Eq. 38-3, or

$$\omega = \sqrt{\frac{1}{LC}} = \sqrt{\frac{1}{(10 \times 10^{-3}\ \text{henry})(1.0 \times 10^{-6}\ \text{farad})}} = 1.0 \times 10^4\ \text{radians/sec.}$$

The time t is then

$$= \frac{\pi}{4\omega} = \frac{\pi}{(4)(1.0 \times 10^4 \text{ radians/sec})} = 7.9 \times 10^{-5} \text{ sec.}$$

What is the frequency ν in cycles/sec? ◀

The stored electric energy in the LC circuit, using Eq. 38–6, is

$$U_E = \frac{1}{2}\frac{q^2}{C} = \frac{{q_m}^2}{2C}\cos^2(\omega t + \phi), \tag{38–8}$$

and the magnetic energy, using Eq. 38–7, is

$$U_B = \tfrac{1}{2}Li^2 = \tfrac{1}{2}L\omega^2{q_m}^2\sin^2(\omega t + \phi).$$

Substituting the expression for ω (Eq. 38–3) into this last equation yields

$$U_B = \frac{{q_m}^2}{2C}\sin^2(\omega t + \phi). \tag{38–9}$$

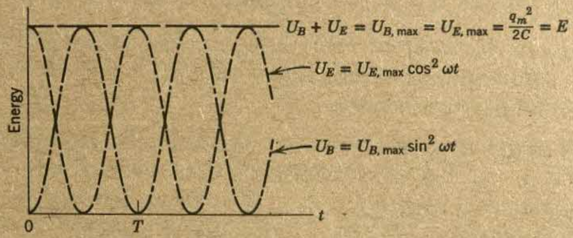


Fig. 38–4 The stored magnetic and electric energy in the circuit of Fig. 38–1. Note that their sum is a constant. The letters indicate corresponding phases of oscillation in Fig. 38–1.

Figure 38–4 shows plots of $U_E(t)$ and $U_B(t)$ for the case of $\phi = 0$. Note that (*a*) the maximum values of U_E and U_B are the same ($= {q_m}^2/2C$); (*b*) at any instant the sum of U_E and U_B is a constant ($= {q_m}^2/2C$); and (*c*) when U_E has its maximum value, U_B is zero and conversely. This analysis supports the qualitative analysis of Section 38–1. The student should compare this discussion with that given in Section 15–4 for the energy transfers in a mass-spring system.

▶ **Example 3.** *The LCR circuit.* (*a*) Derive an expression for the quantity $q(t)$ for a single-loop circuit containing a resistance R as well as an inductance L and a capacitance C. (*b*) After how long a time will the charge oscillations decay to half-amplitude if $L = 10$ mh, $C = 1.0\ \mu$f, and $R = 0.1$ ohm?

(*a*) If U is the total stored field energy, we have, as before,

$$U = \tfrac{1}{2}Li^2 + \frac{1}{2}\frac{q^2}{C}.$$

U is no longer constant but rather

$$\frac{dU}{dt} = -i^2R,$$

the minus sign signifying that the stored energy U *decreases* with time, being converted to Joule heat at the rate i^2R. Combining these two equations leads to

$$Li\frac{di}{dt} + \frac{q}{C}\frac{dq}{dt} = -i^2R.$$

Substituting dq/dt for i and d^2q/dt^2 for di/dt leads finally to

$$L\frac{d^2q}{dt^2} + R\frac{dq}{dt} + \frac{1}{C}q = 0,$$

which is the differential equation that describes the damped LC oscillations. If we put $R = 0$, this equation reduces, as expected, to Eq. 38-5.

The student should compare this differential equation for damped LC oscillations with Eq. 15-37, or

$$m\frac{d^2x}{dt^2} + b\frac{dx}{dt} + kx = 0, \qquad (15\text{-}28)$$

which describes damped mass-spring oscillations. Once again the equations are mathematically identical, the resistance R corresponding to the mechanical damping constant b.

The solution of the LCR circuit follows at once, by correspondence, from the solution of Eq. 15-37. It is (see Eqs. 15-38 and 15-39) for R reasonably small, and for an initial condition in which the capacitor has a maximum charge

$$q = q_m e^{-Rt/2L}\cos\omega' t, \qquad (38\text{-}10)$$

where

$$\omega' = \sqrt{\frac{1}{LC} - \left(\frac{R}{2L}\right)^2}. \qquad (38\text{-}11)$$

Note that Eq. 38-10, which can be described as a cosine function with an exponentially decreasing amplitude, is the equation of the decay curve of Fig. 38-3. Note, too (Eq. 38-11), that the presence of resistance reduces the oscillation frequency. These two equations reduce to familiar results as $R \to 0$.

(*b*) The oscillation amplitude will have decreased to half when the amplitude factor $e^{-Rt/2L}$ in Eq. 38-10 has the value one-half, or

$$\tfrac{1}{2} = e^{-Rt/2L},$$

which leads readily to

$$t = \frac{2L}{R}\ln 2 = \frac{(2)(10 \times 10^{-3}\text{ henry})(0.69)}{0.10\text{ ohm}} = 0.14\text{ sec}.$$

The angular frequency, from Eq. 38-11, is

$$\omega' = \sqrt{\frac{1}{(10 \times 10^{-3}\text{ henry})(1.0 \times 10^{-6}\text{ farad})} - \left(\frac{0.10\text{ ohm}}{2 \times 10 \times 10^{-3}\text{ henry}}\right)^2}$$
$$= \sqrt{10^8\text{ radians/sec}^2 - 25\text{ radians/sec}^2} = 1.0 \times 10^4\text{ radians/sec}.$$

Note that the second term is rather small, so that in this case, as in many practical cases, the resistance has a negligible effect on the frequency. The student should show that 0.14 sec, the time at which the oscillations decrease to half-amplitude, corresponds to about 220 cycles of oscillation. The damping is much less severe than that illustrated in Fig. 38-3. ◀

38–4 Forced Oscillations and Resonance

Figure 38–5 shows an LCR circuit containing a sinusoidally varying emf $\mathcal{E}(t)$ given by

$$\mathcal{E} = \mathcal{E}_m \cos \omega'' t,$$

in which ω'' can be varied at will. The emf $\mathcal{E}(t)$ might be provided by a variable-frequency oscillator. What amplitudes of electromagnetic oscillations are set up in this circuit for various angular frequencies ω'' of the "driving force"?

The problem corresponds to that of forced oscillations in the damped mass-spring system of Section 15–10. The differential equation describing that motion is

$$m \frac{d^2x}{dt^2} + b \frac{dx}{dt} + kx = F_m \cos \omega'' t, \tag{15–31}$$

where ω'' is the angular frequency of the external periodic driving force applied to the system and F_m is its amplitude.

For the circuit of Fig. 38–5, the differential equation that follows from the correspondences of p. 947 and the additional reasonable correspondence of $\mathcal{E}$ to F, is

$$L \frac{d^2q}{dt^2} + R \frac{dq}{dt} + \frac{1}{C} q = \mathcal{E}_m \cos \omega'' t. \tag{38–12}$$

This equation can also be derived by applying the energy conservation principle to the circuit of Fig. 38–5.

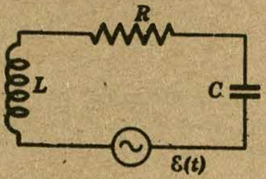


Fig. 38–5 An LCR circuit containing a sinusoidal emf of angular frequency ω''.

We can use the correspondences of p. 947 to write down the solution to Eq. 38–12. Starting from Eq. 15–4 and making the appropriate substitutions, we obtain

$$q = \frac{\mathcal{E}_m}{G} \sin (\omega'' t - \phi), \tag{38–13}$$

where

$$G = \sqrt{\left(\omega''^2 L - \frac{1}{C}\right)^2 + R^2 \omega''^2},$$

and ϕ, the *phase angle* between the "driving force" and the "response," is given by

$$\phi = \cos^{-1} \frac{R\omega''}{G}.$$

As often as not, we are interested in the current $i(t)$ in the circuit, rather than the charge; the current corresponds to the velocity $v(t)$ of the moving mass in Section 15–10. We can find $i(t)$ by differentiating Eq. 38–13, or

$$i = \frac{dq}{dt} = \frac{\omega'' \mathcal{E}_m}{G} \cos (\omega'' t - \alpha) = i_m \cos (\omega'' t - \alpha).$$

The amplitude i_m of the current oscillations is given, from these equations, by

$$i_m = \frac{\omega''\mathcal{E}_m}{G} = \frac{\mathcal{E}_m}{\sqrt{\left(\omega''L - \frac{1}{\omega''C}\right)^2 + R^2}}. \tag{38-14}$$

Inspection of Eq. 38-14 shows that the current (not the charge; see Question 5) will have its maximum amplitude when

$$\omega''L = \frac{1}{\omega''C},$$

which can be written as

$$\omega'' = \sqrt{\frac{1}{LC}}. \tag{38-15}$$

Comparison with Eq. 38-3 shows that the maximum amplitude of the current oscillations occurs when the frequency ω'' of the applied emf is exactly equal to the natural (undamped) frequency ω of the system.

At *resonance* ($\omega'' = \omega$) the amplitude of the current oscillations is determined entirely by the resistance; this follows by combining Eqs. 38-14 and 38-15, or

$$i_m = \frac{\mathcal{E}_m}{R} \qquad \text{(at resonance)}.$$

Figure 38-6 shows i_m as a function of ω'' for an oscillating *LCR* circuit containing three different values of resistance. Note that the smaller the resistance, the sharper the resonance curve. The sharpness of such a curve is measured by its ***half-width***, which is the difference between two frequencies, each of which corresponds to a current amplitude of one-half the maximum current amplitude. The half-width of the curve for R = 10 ohms in Fig. 38-6 is shown in that figure by the arrow marked $\Delta\omega$.

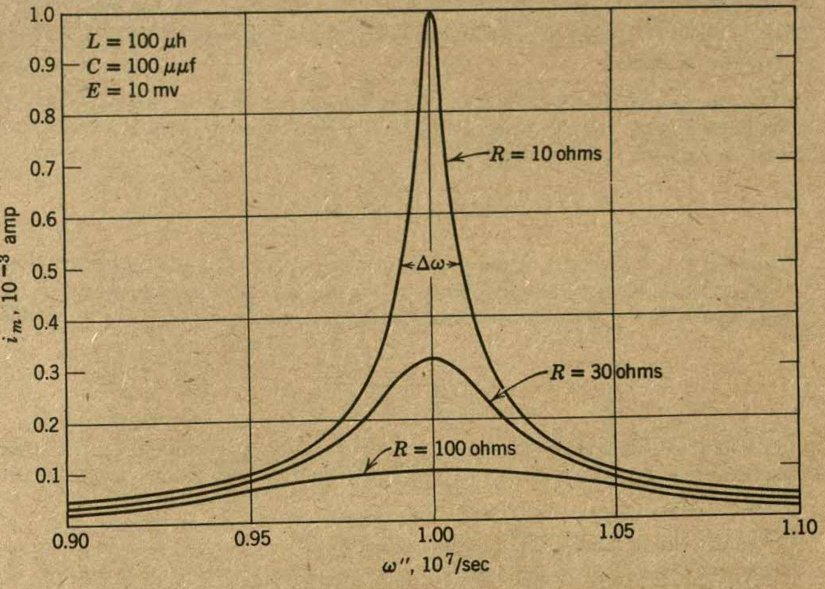


Fig. 38-6 The current amplitude as a function of frequency in the circuit of Fig. 38-5. The arrow marked $\Delta\omega$ on the curve for R = 10 ohms is the *half-width* of that curve.

38-5 Lumped and Distributed Elements

In the oscillating mass-spring system the two kinds of energy involved appear in separate parts of the system, the potential energy being stored in the spring and the kinetic energy in the moving mass. An *acoustic cavity resonator*, such as an organ pipe, is a mechanical oscillating system in which the two forms of energy are *not* separated in space. Kinetic energy, associated with moving air in the cavity, and potential energy, associated with compressions or rarefactions of the air, can both be present throughout the volume of the cavity. Such a cavity is an example of an oscillating system with *distributed* rather than *lumped* (as in the mass-spring system) elements.

A similar distinction exists in electromagnetic systems. The LC circuit of Fig. 38–1 is an example of lumped elements, in that the two kinds of energy are stored in rather different places; the circuit is described by giving the (lumped) system parameters L and C. Actually, in modern engineering practice and in physics research electromagnetic systems with *distributed elements* play a fundamental role.

Figure 38–7, a series of "snapshots" taken one-eighth of a cycle apart, shows the pressure and velocity variations in the fundamental mode of a particular acoustic resonator. There is a pressure node at the center and a pressure antinode at each end. There is a velocity * node at each end and a velocity antinode at the center. When the pressure variation is the greatest, the velocity is zero (Figs. 38–7*a* and *e*). When the pressure is uniform, the velocities have their maximum values (Figs. 38–7*c*, and *g*).

The energy in the acoustic resonator alternates between kinetic energy associated with the moving gas and potential energy associated with the compression and rarefaction of the gas. In Figs. 38–7*c* and *g* the energy is all kinetic and in Figs. 38–7*a* and *e* it is all potential. In intermediate phases it is part of each.

The kinetic energy of a small mass Δm of the gas, which is moving parallel to the cylinder axis with a speed v_g, is $\frac{1}{2}\Delta m v_g^2$. The *kinetic energy density*, that is, the kinetic energy per unit volume, is

$$u_K = \frac{1}{2}\frac{\Delta m}{\Delta V}v_g^2 = \tfrac{1}{2}\rho_0 v_g^2,$$

in which ΔV is the volume of the gas element and ρ_0 is the mean density of the gas.

The potential energy per unit volume in the gas, that is, the *potential energy density*, associated with the compressions and rarefactions of the gas may be given by

$$u_P = \tfrac{1}{2}B\left(\frac{\Delta\rho}{\rho_0}\right)^2.$$

Here B is the bulk modulus of elasticity of the gas and $\Delta\rho/\rho_0$, which is positive for a compression and negative for a rarefaction, is the fractional change in gas density.

---

* The velocity of interest here is the directed velocity v_g of small volume elements of the gas which are, however, large enough to contain a great number of molecules. The thermal velocities of the molecules have no directional preference and are ignored.

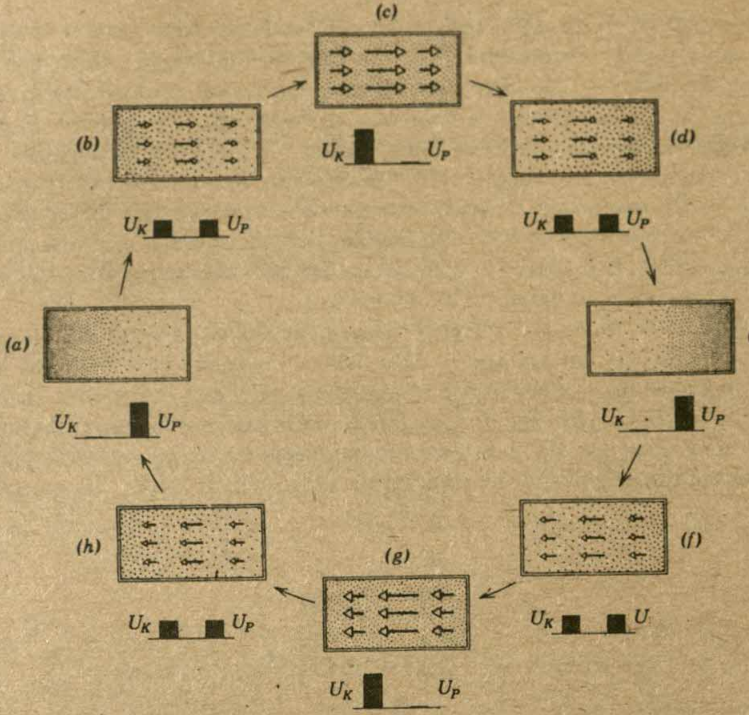


Fig. 38-7 Showing eight stages in a cycle of oscillation of a cylindrical acoustic resonant cavity. The bar graphs below each figure show the kinetic and potential energy. The arrows represent the directed velocities of small volume elements of the gas. Compare with Fig. 38-1.

The angular frequency of oscillation for the cavity of Fig. 38-7, in the fundamental (or lowest frequency) mode shown in that figure, is found from

$$\omega_1 = 2\pi\nu = \frac{2\pi v}{\lambda} = \frac{\pi v}{l},$$

where v is the speed of sound in the gas and l is the length of the cavity. From Eq. 20-1 we may write v as $\sqrt{B_0/\rho_0}$. Note that in the above we have put $\lambda = 2l$, corresponding to the fundamental mode. What are the frequencies ω_2, ω_3, etc., of the higher frequency modes?

38-6 Electromagnetic Cavity Oscillator

Consider now a second closed cylinder, of radius a and length l, whose walls are made of copper or some other good conductor. A system of oscillating electric and magnetic fields can be set up in such a cavity even if, as in the common case that we consider, the cavity contains no material medium. Such an *electromagnetic cavity resonator* is a distributed electromagnetic oscillator, in contrast to an LC circuit, which is a lumped system. As for the acoustic resonator, many modes of oscillation with discrete frequencies are possible; we describe only the fundamental mode, which has the lowest frequency. The cavity oscillations can be set up by suitably connecting the cavity, through a small hole in its side wall, to a source of electromagnetic radiation such as a magnetron. If the cavity dimensions are of the order of a few centimeters, the resonant frequencies will be of the order of 10^{10} cycles/sec. This is far higher than the acoustic frequencies in cavities of the same size, reflecting the fact that the speed of electromagnetic disturbances in free space ($= 3 \times 10^8$ meters/sec) is much greater than the speed of sound in air (about 350 meters/sec).

Figure 38–8, which is a series of "snapshots" taken one-eighth of a cycle apart, shows, by the horizontal lines, how the electric field **E** varies with time in the cavity. The electric lines of force originate on positive charges at one end of the cylinder and terminate on negative charges at the other end. As **E** changes with time, eventually reversing itself, currents flow from one end of the cylinder to the other on the inner cylinder wall. At any point

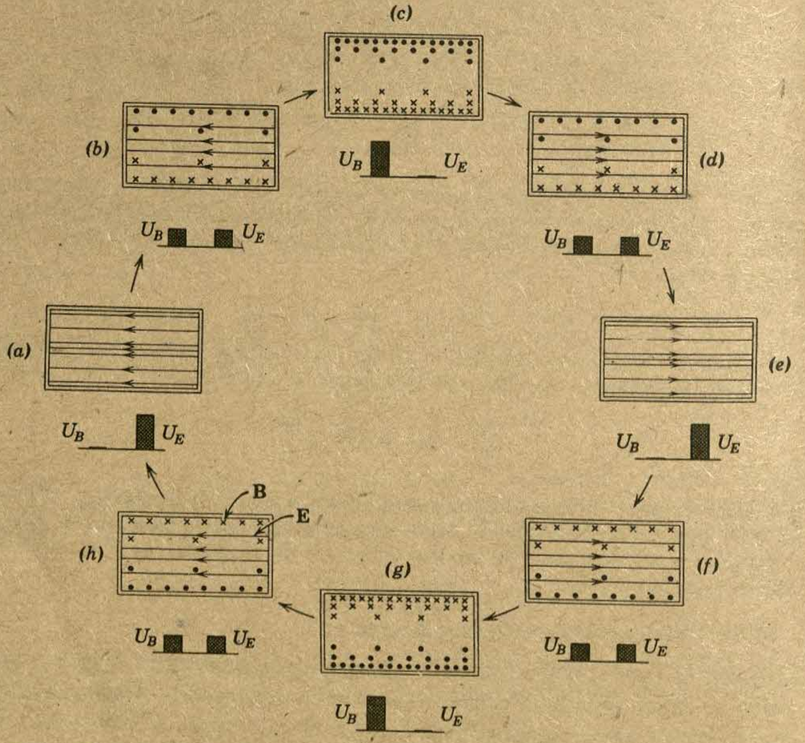


Fig. 38–8 Showing eight stages in a cycle of oscillation of a cylindrical electromagnetic resonant cavity. The bar graphs below each figure show the stored electric and magnetic energy. The dots and crosses represent circular lines of **B**; the horizontal lines represent **E**. Compare with Fig. 38–7.

in the cavity, energy is stored in the electric field in an amount per unit volume given by Eq. 30–27, or

$$u_E = \tfrac{1}{2}\epsilon_0 E^2. \tag{38–16}$$

Figure 38–8 also shows, by the dots and crosses, how the magnetic field **B** varies with time. The magnetic lines form circles about the cylinder axis. Note that the magnetic field has a maximum value when the electric field is zero, and conversely. At any point in the cavity energy is stored in the magnetic field in an amount per unit volume given by Eq. 36–19, or

$$u_B = \frac{1}{2\mu_0} B^2. \tag{38–17}$$

Thus, as in the *LC* circuit, energy is shuttled back and forth between the electric and the magnetic fields. However, these fields no longer occupy completely separate regions of space.

We state without proof that the angular frequency of oscillation for the electromagnetic cavity of Fig. 38–8 is, in the fundamental mode shown in that figure,

$$\omega_1 = \frac{1.19c}{a},$$

in which a is the cavity radius and c is the speed of electromagnetic radiations in free space. We will see in Section 39–5 that c may be written in terms of electromagnetic quantities as $1/\sqrt{\mu_0\epsilon_0}$. As the field patterns of Fig. 38–8 suggest, the resonant frequency of oscillation of the cavity, for the mode of oscillation shown, depends only on the radius of the cavity and not on its length.

Table 38–2 summarizes some characteristics of the four oscillating systems that we have discussed so far. For lumped systems it gives expressions for the two kinds

Table 38–2

FOUR OSCILLATING SYSTEMS

	Mechanical Systems	Electromagnetic Systems
Lumped systems	Mass + spring $U_K = \frac{1}{2}mv^2$ $U_P = \frac{1}{2}kx^2$ $\omega = \sqrt{\frac{k}{m}}$	*LC* circuit $U_B = \frac{1}{2}Li^2$ $U_E = \frac{1}{2}(1/C)q^2$ $\omega = \sqrt{\frac{(1/C)}{L}}$
Distributed systems	Acoustic cavity $u_K = \frac{1}{2}\rho_0 {v_g}^2$ $u_P = \frac{1}{2}B(\Delta\rho/\rho_0)^2$ $\omega_1 = \frac{3.14v}{l}; \quad v = \sqrt{\frac{B}{\rho_0}}$	Electromagnetic cavity $u_B = \frac{1}{2}(1/\mu_0)B^2$ $u_E = \frac{1}{2}\epsilon_0 E^2$ $\omega_1 = \frac{1.19c}{a}; \quad c = \sqrt{\frac{1}{\epsilon_0\mu_0}}$

of energy involved and for the (single) oscillation frequency. For the distributed systems it gives expressions for the two kinds of energy density involved and for the oscillation frequency in the fundamental mode. The student should study carefully all the correspondences, similarities, and differences that occur in the table.

The expression given above for the fundamental resonant angular frequency ω_1 of the cavity is derived by applying the basic equations of electromagnetism (Table 38–3) to the space inside the cavity and by invoking this boundary condition: **E** *must be zero inside the cavity wall and can have no tangential component anywhere on the cavity wall.* If this were not true, an infinite current would be set up in the assumed resistanceless wall by the tangential component of **E**. This boundary condition is similar in spirit to the requirement that a clamped string must have zero amplitude of oscillation at the points at which it is clamped or that a velocity node must exist at the end walls of an acoustic cavity resonator. Granted this boundary condition, on **E**, it can be shown that, assuming a perfectly conducting wall, (*a*) no time-varying magnetic field can exist inside the cavity wall and (*b*) no time-varying currents can exist inside the wall. A tangential magnetic field can exist *on the surface*, however; surface charges can exist, and surface currents can flow.

▶ **Example 4.** In the cavity of Fig. 38–8, what is the relationship between the "average" value of **E** throughout the cavity, measured at the instant corresponding to Fig. 38–8*a*, to the "average" value of **B**, measured at the instant corresponding to Fig. 38–8*c*?

At the first instant the energy is all electric and at the second it is all magnetic. The total energy U, found by integrating the energy density over the volume of the cavity, must be the same at these two instants, or

$$U = \int u_{E,m}\,dV = \int u_{B,m}\,dV,$$

where dV is a volume element in the cavity and $u_{E,m}$ and $u_{B,m}$ are the *maximum* values of u_E and of u_B at the site of this volume element; these maximum values occur one-fourth of a cycle apart, as Fig. 38–8 shows. Substituting Eqs. 38–16 and 38–17 leads to

$$\int \frac{\epsilon_0 E_m{}^2}{2}\,dV = \int \frac{B_m{}^2}{2\mu_0}\,dV$$

or

$$\mu_0\epsilon_0 \int E_m{}^2\,dV = \int B_m{}^2\,dV.$$

The quantity $\int E_m{}^2\,dV$ can be written as $\overline{E_m{}^2}V$, where V is the cavity volume and $\overline{E_m{}^2}$ is the average value of $E_m{}^2$ throughout the cavity. Treating B_m in the same way leads to

$$\mu_0\epsilon_0\overline{E_m{}^2} = \overline{B_m{}^2}$$

or, taking square roots,

$$\sqrt{\overline{B_m{}^2}} = \sqrt{\mu_0\epsilon_0}\,\sqrt{\overline{E_m{}^2}}.$$

We can represent $\sqrt{\overline{B_m{}^2}}$ by B_{rms}, a so-called "root-mean-square" value of B_m. In computing B_{rms}, note that the averaging is done throughout the volume of the cavity, at the instant corresponding to Fig. 38–8*c*. It is not a time average for a particular point in the cavity. Doing the same for **E** yields

$$\begin{aligned} B_{\text{rms}} &= \sqrt{\mu_0\epsilon_0}\,E_{\text{rms}} = \sqrt{(4\pi \times 10^{-7}\ \text{weber/amp-m})(8.9 \times 10^{-12}\ \text{coul}^2/\text{nt-m}^2)}\,E_{\text{rms}} \\ &= (3.3 \times 10^{-9}\ \text{sec/meter})E_{\text{rms}}. \end{aligned}$$

If E_{rms} equals 10^4 volts/meter, a reasonable value, then

$$B_{\text{rms}} = (3.3 \times 10^{-9}\ \text{sec/meter})(10^4\ \text{volts/meter})$$

$$= 3.3 \times 10^{-5}\ \text{weber/meter}^2 = 0.33\ \text{gauss}.$$

What is the total stored energy in the cavity under these conditions, assuming the cavity to be 10 cm long and 3.0 cm in diameter?

The student will recall that in Example 6, Chapter 36, we showed the energy density for a magnetic field of "ordinary" laboratory magnitude (say, 1 weber/meter^2) to be enormously greater than that for an electric field of "ordinary" magnitude (say, 10^5 volts/meter). This fact is consistent with the present example. ◀

38-7 Induced Magnetic Fields

To understand the electromagnetic cavity oscillations in terms of electromagnetic theory, we must complete our description of the basic equations of electromagnetism by introducing a new concept, namely, that *a changing electric field produces a magnetic field.* This concept, which is the symmetrical counterpart of Faraday's law of induction, is of fundamental significance. We will develop this concept by a symmetry argument and will let the agreement with experiment of our final conclusions speak for itself. This comparison with experiment, which is worked out largely in Chapters 39 and 40, forms one of the chief experimental bases of electromagnetic theory. A central achievement was the demonstration that the experimentally measured speed c of visible light in free space could be related to purely electromagnetic quantities by

$$c = \frac{1}{\sqrt{\mu_0 \epsilon_0}}. \tag{38-18}$$

This demonstration not only revealed optics as a branch of electromagnetism but led directly to the concept of the electromagnetic spectrum, which in turn resulted in the discovery of radio waves.

Figure 38-9*a* shows a uniform electric field **E** filling a cylindrical region of space. It might be produced by a circular parallel-plate capacitor, as suggested in Fig. 38-9*b*. We assume that E is increasing at a steady rate dE/dt, which means that charge must be supplied to the capacitor plates at a steady rate; to supply this charge requires a steady current i into the positive plate and an equal steady current i out of the negative plate.

If a sufficiently delicate experiment could be performed, it would be found that *a magnetic field is set up by this changing electric field.* Figure 38-9*a* shows **B** for four arbitrary points. Figure 38-9 suggests a beautiful example of the symmetry of nature. A changing *magnetic* field induces an *electric* field (Faraday's law); now we see that a changing *electric* field induces a *magnetic* field.

To describe this new effect quantitatively, we are guided by analogy with Faraday's law of induction,

$$\oint \mathbf{E} \cdot d\mathbf{l} = -\frac{d\Phi_B}{dt}, \tag{38-19}$$

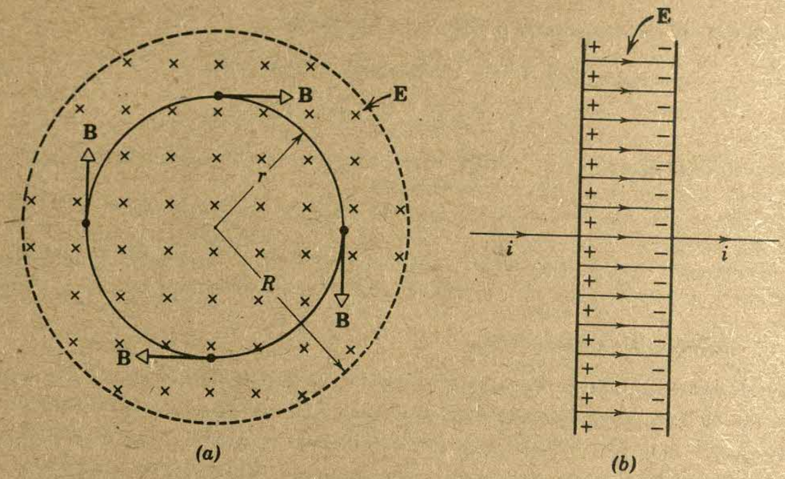


Fig. 38–9 (*a*) Showing the induced magnetic fields **B** at four points, produced by a changing electric field **E**. The electric field is increasing in magnitude. Compare Fig. 35–10. (*b*) Such a changing electric field may be produced by charging a parallel-plate capacitor as shown.

which asserts that an electric field (left term) is produced by a changing magnetic field (right term). For the symmetrical counterpart we might write *

$$\oint \mathbf{B}\cdot d\mathbf{l} = \mu_0\epsilon_0 \frac{d\Phi_E}{dt}. \tag{38–20}$$

Equation 38–20 asserts that a magnetic field (left term) can be produced by a changing electric field (right term). The student should compare carefully Fig. 35–10, which illustrates the production of an electric field by a changing magnetic field, with Fig. 38–9*a*. In each case the appropriate flux Φ_B or Φ_E is *increasing*. However, experiment shows that the lines of **E** in Fig. 35–10 are *counterclockwise*, whereas those of **B** in Fig. 38–9*a* are *clockwise*. This difference requires that the minus sign of Eq. 38–19 be omitted from Eq. 38–20.

In Section 34–1 we saw that a magnetic field can also be set up by a current in a wire. We described this quantitatively by Ampère's law:

$$\oint \mathbf{B}\cdot d\mathbf{l} = \mu_0 i,$$

in which i is the conduction current passing through the loop around which the line integral is taken. Thus there are at least two ways of setting up a magnetic field: (*a*) by a changing electric field and (*b*) by a current. In general, both possibilities must be allowed for, or *

---

* Our system of units requires that we insert the constants ϵ_0 and μ_0 in Eq. 38–20. In some unit systems they would not appear.

$$\oint \mathbf{B}\cdot d\mathbf{l} = \mu_0\epsilon_0 \frac{d\Phi_E}{dt} + \mu_0 i. \tag{38-21}$$

Maxwell is responsible for this important generalization of Ampère's law.

In Chapter 34 we assumed that no changing electric fields were present so that the term $d\Phi_E/dt$ in Eq. 38-21 is zero. In the discussion just given we assumed that there were no conduction currents in the space containing the electric field. Thus the term i in Eq. 38-21 is zero. We see now that each of these situations is a special case.

▶ **Example 5.** A parallel-plate capacitor with circular plates is being charged as in Fig. 38-9. (*a*) Derive an expression for the induced magnetic field at various radii r. Consider both $r < R$ and $r > R$.

From Eq. 38-20,

$$\oint \mathbf{B}\cdot d\mathbf{l} = \mu_0\epsilon_0 \frac{d\Phi_E}{dt},$$

we can write, for $r \leqq R$,

$$(B)(2\pi r) = \mu_0\epsilon_0 \frac{d}{dt}[(E)(\pi r^2)] = \mu_0\epsilon_0\pi r^2 \frac{dE}{dt}.$$

Solving for B yields $$B = \tfrac{1}{2}\mu_0\epsilon_0 r \frac{dE}{dt} \qquad (r \leqq R).$$

For $r \geqq R$, Eq. 38-20 yields

$$(B)(2\pi r) = \mu_0\epsilon_0 \frac{d}{dt}[(E)(\pi R^2)] = \mu_0\epsilon_0\pi R^2 \left(\frac{dE}{dt}\right),$$

or $$B = \frac{\mu_0\epsilon_0 R^2}{2r}\frac{dE}{dt} \qquad (r \geqq R).$$

(*b*) Find B at $r = R$ for $dE/dt = 10^{12}$ volts/m-sec and for $R = 5.0$ cm. At $r = R$ the two equations for B reduce to the same expression, or

$$\begin{aligned} B &= \tfrac{1}{2}\mu_0\epsilon_0 R \frac{dE}{dt} \\ &= (\tfrac{1}{2})(4\pi \times 10^{-7}\text{ weber/amp-m})(8.9 \times 10^{-12}\text{ coul}^2/\text{nt-m}^2) \\ &\quad (5.0 \times 10^{-2}\text{ meter})(10^{12}\text{ volts/m-sec}) \\ &= 2.8 \times 10^{-7}\text{ weber/meter}^2 = 0.0028\text{ gauss.} \end{aligned}$$

This shows that the induced magnetic fields in this example are so small that they can scarcely be measured with simple apparatus, in sharp contrast to induced *electric* fields (Faraday's law), which can be demonstrated easily. This experimental dif-

---

* Actually, there is a third way of setting up a magnetic field, by the use of magnetized bodies. In Section 37-7 we saw that this could be accounted for by inserting a *magnetizing current* term i_M on the right side of Ampère's law. This law would then read, in its full generality,

$$\oint \mathbf{B}\cdot d\mathbf{l} = \mu_0\epsilon_0 \frac{d\Phi_E}{dt} + \mu_0 i + \mu_0 i_M.$$

In all that follows we assume that no magnetic materials are present so that $i_M = 0$.

ference is in part due to the fact that induced emfs can easily be multiplied by using a coil of many turns. No technique of comparable efficiency exists for magnetic fields. In experiments involving oscillations at very high frequencies dE/dt above can be very large, resulting in significantly larger values of the induced magnetic field. ◀

38–8 Displacement Current

Equation 38–21 shows that the term $\epsilon_0\, d\Phi_E/dt$ has the dimensions of a current. Even though no motion of charge is involved, there are advantages in giving this term the name *displacement* * *current.* Thus we can say that a magnetic field can be set up either by a conduction current i or by a displacement current i_d $(= \epsilon_0\, d\Phi_E/dt)$, and Eq. 38–21 can be rewritten as †

$$\oint \mathbf{B}\cdot d\mathbf{l} = \mu_0(i_d + i). \qquad (38\text{–}22)$$

The concept of displacement current permits us to retain the notion that *current is continuous,* a principle established for steady conduction currents in Section 31–1. In Fig. 38–9*b*, for example, a current i enters the positive plate and leaves the negative plate. The *conduction* current is *not* continuous across the capacitor gap because no charge is transported across this gap. However, the displacement current i_d in the gap will prove to be exactly i, thus retaining the concept of the continuity of current.

To calculate the displacement current, recall (see Eq. 30–5) that E in the gap is given by

$$E = \frac{q}{\epsilon_0 A}.$$

Differentiating gives
$$\frac{dE}{dt} = \frac{1}{\epsilon_0 A}\frac{dq}{dt} = \frac{1}{\epsilon_0 A}\,i.$$

The displacement current i_d is by definition

$$i_d = \epsilon_0 \frac{d\Phi_E}{dt} = \epsilon_0 \frac{d(EA)}{dt} = \epsilon_0 A\, \frac{dE}{dt}.$$

Combining these two equations leads to

$$i_d = (\epsilon_0 A)\left(\frac{1}{\epsilon_0 A}\,i\right) = i,$$

which shows that the displacement current in the gap is identical with the conduction current in the lead wires.

---

* The word "displacement" was introduced for historical reasons that need not concern us here.

† We may write this more generally, taking the presence of magnetic materials into account, as

$$\oint \mathbf{B}\cdot d\mathbf{l} = \mu_0(i_d + i + i_M).$$

See the footnote on p. 961.

▶ **Example 6.** What is the displacement current for the capacitor of Example 5? From the definition of displacement current,

$$\begin{aligned} i_d &= \epsilon_0 \frac{d\Phi_E}{dt} = \epsilon_0 \frac{d}{dt}[(E)(\pi R^2)] = \epsilon_0 \pi R^2 \frac{dE}{dt} \\ &= (8.9 \times 10^{-12}\,\text{coul}^2/\text{nt-m}^2)(\pi)(5.0 \times 10^{-2}\,\text{meter})^2(10^{12}\,\text{volts/m-sec}) \\ &= 0.070\,\text{amp}. \end{aligned}$$

Even though this displacement current is reasonably large, it produces only a small magnetic field (see Example 5) because it is spread out over a large area. ◀

38–9 Maxwell's Equations

Equation 38–21 completes our presentation of the basic equations of electromagnetism, called *Maxwell's equations.* They are summarized in Table 38–3. All equations of physics that serve, as these do, to correlate experiments in a vast area and to predict new results have a certain beauty about them and can be appreciated, by those who understand them, on an aesthetic level. This is true for Newton's laws of motion, for the laws of thermodynamics, for the theory of relativity, and for the theories of quantum physics. As for Maxwell's equations, the German physicist Ludwig Boltzmann (quoting a line from Goethe) wrote "Was it a god who wrote these lines. . . ." In more recent times J. R. Pierce,* in a book chapter entitled "Maxwell's Wonderful Equations" says: "To anyone who is motivated by anything beyond the most narrowly practical, it is worth while to understand Maxwell's equations simply for the good of his soul." The scope of these equations is remarkable, including as it does the fundamental operating principles of all large-scale electromagnetic devices such as motors, cyclotrons, electronic computers, television, and microwave radar.

38–10 Maxwell's Equations and Cavity Oscillations

In this section we show how the oscillations of an electromagnetic cavity can be understood in terms of Maxwell's equations. A completely formal treatment, which is beyond our scope here, would start from these equations and would end with mathematical expressions for the variation of **B** and **E** with time and with position in the cavity for all modes of oscillation of the cavity. We confine ourselves to the fundamental mode only, illustrated in Fig. 38–8, for which we *postulated* the variations of **B** and **E** given in that figure; we will show that these postulated fields are completely consistent with Maxwell's equations.

Figure 38–10 presents two views of the cavity of Fig. 38–8*d*, in which both electric and magnetic fields are present. Study of Fig. 38–8 reveals that **B** is *decreasing* in magnitude and **E** is *increasing* at this phase of the cycle of oscillation. Let us apply Faraday's law,

$$\oint \mathbf{E} \cdot d\mathbf{l} = -\frac{d\Phi_B}{dt},$$

---

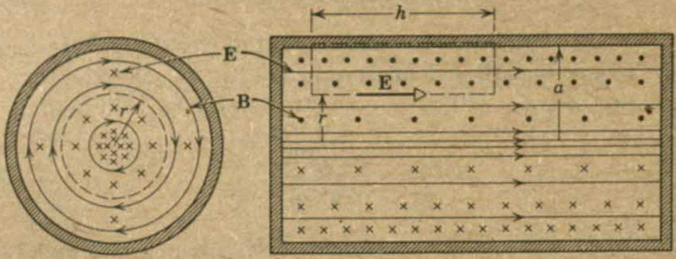
* *Electrons, Waves and Messages*, Hanover House, 1956. This book is recommended as collateral reading in electromagnetism.

Table 38–3

THE BASIC EQUATIONS OF ELECTROMAGNETISM (MAXWELL'S EQUATIONS) *

Name	Equation	Describes	Crucial Experiment	Text Reference
Gauss's law for electricity	$\epsilon_0 \oint \mathbf{E}\cdot d\mathbf{S} = q$	Charge and the electric field	1. Like charges repel and unlike charges attract, as the inverse square of their separation. 1′. A charge on an insulated conductor moves to its outer surface.	Chapter 28
Gauss's law for magnetism	$\oint \mathbf{B}\cdot d\mathbf{S} = 0$	The magnetic field	2. It is impossible to create an isolated magnetic pole.	Section 37–2
Ampère's law (as extended by Maxwell)	$\oint \mathbf{B}\cdot d\mathbf{l} = \mu_0\left(\epsilon_0 \frac{d\Phi_E}{dt} + i\right)$	The magnetic effect of a changing electric field or of a current	3. The speed of light can be calculated from purely electromagnetic measurements. 3′. A current in a wire sets up a magnetic field near the wire.	Section 39–5 Chapter 34
Faraday's law of induction	$\oint \mathbf{E}\cdot d\mathbf{l} = -\frac{d\Phi_B}{dt}$	The electrical effect of a changing magnetic field	4. A bar magnet, thrust through a closed loop of wire, will set up a current in the loop.	Chapter 35

* Written on the assumption that no dielectric or magnetic material is present.



(a) (b)

Fig. 38-10 Two cross sections of an electromagnetic resonant cavity at a phase of oscillation corresponding to Fig. 38-8*d*. (*a*) The dashed circle is a path suitable for applying Ampère's law. (*b*) The dashed rectangle is a path suitable for applying Faraday's law.

to the rectangle of dimensions h and $a - r$. There is a definite flux Φ_B through the rectangular path in question, and this flux is decreasing with time because **B** is decreasing. The line integral above is

$$\oint \mathbf{E} \cdot d\mathbf{l} = hE(r),$$

in which $E(r)$ is the value of E at a radius r from the center of the cavity.

Note that **E** equals zero for the upper leg of the integration path (which lies in the cavity wall) and that **E** and $d\mathbf{l}$ are at right angles on the two side legs. Combining these equations yields

$$E(r) = -\frac{1}{h}\frac{d\Phi_B}{dt}. \tag{38-23}$$

Equation 38-23 shows that $E(r)$ depends on the rate at which Φ_B through the path shown is changing with time and that it has its maximum value when $d\Phi_B/dt$ is a maximum. This occurs when **B** is zero, that is, when **B** is changing its direction; the student will recall that a sine or cosine is changing most rapidly, that is, it has the steepest tangent, at the instant it crosses the axis between positive and negative values. Thus the electric field pattern in the cavity will have its *maximum* value when the magnetic field is *zero* everywhere, consistent with Figs. 38-8*a* and *e* and with the concept of the interchange of energy between electric and magnetic forms. The student should demonstrate, by applying Lenz's law, that the electric field in Fig. 38-10*b* points to the right as shown, if the magnetic field is decreasing.

Figure 38-10*a* shows an end view of the cavity; the electric lines of force are entering the page at right angles to the page and the magnetic lines form clockwise circles. Let us apply Ampère's law in the form

$$\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0\epsilon_0 \frac{d\Phi_E}{dt} + \mu_0 i, \tag{38-21}$$

to the circular path of radius r shown in the figure. No charge is transported through the ring so that the conduction current i in Eq. 38–21 is zero. The line integral on the left is $(B)(2\pi r)$ so that the equation reduces to

$$B(r) = \frac{\mu_0 \epsilon_0}{2\pi r}\frac{d\Phi_E}{dt}. \tag{38–24}$$

Equation 38–24 shows that the magnetic field $B(r)$ is proportional to the rate at which the electric flux Φ_E through the ring is changing with time. The field $B(r)$ has its maximum value when $d\Phi_E/dt$ is at its maximum; this occurs when $\mathbf{E} = 0$, that is, when **E** is reversing its direction. Thus we see that **B** has its *maximum* value when **E** is *zero* for all points in the cavity. This is consistent with Figs. 38–8*c* and *g* and with the concept of the interchange of energy between electric and magnetic forms. A comparison with Fig. 38–9*a*, which, like Fig. 38–10*a*, corresponds to an increasing electric field, shows that the lines of **B** are indeed clockwise, as viewed along the direction of the electric field.

Comparison of Eqs. 38–23 and 38–24 suggests the complete interdependence of **B** and **E** in the cavity. As the magnetic field changes with time, it induces the electric field in a way described by Faraday's law. The electric field, which also changes with time, induces the magnetic field in a way described by Maxwell's extension of Ampère's law. The oscillations, once established, sustain each other and would continue indefinitely were it not for losses due to Joule heating in the cavity walls or leakage of energy from openings that might be present in the walls. In Chapter 39 we show that this interplay of **B** and **E** occurs not only in standing electromagnetic waves in cavities but also in traveling electromagnetic waves, such as radio waves or visible light.

Let us now analyze the currents—both conduction and displacement—that occur in the cavity and examine their connections to the electric and the magnetic fields. Figure 38–11 shows two views of the cavity, taken at an instant corresponding to that of Fig. 38–10. For simplicity the fields **E** and **B** are not shown; the arrows represent currents.

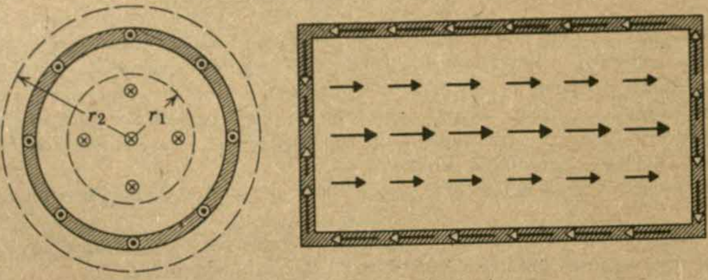
Since **E** is increasing at this instant, the positive charge on the left end cap must be increasing. Thus there must be conduction currents in the walls pointing from right to left in Fig. 38–11*b*. These currents are also shown by the dots (representing the tips of arrows) near the cavity walls in Fig. 38–11*a*.

Bearing in mind that $\epsilon_0\, d\Phi_E/dt$ is a displacement current, we can write Eq. 38–24 as

$$B(r) = \frac{\mu_0}{2\pi r}\epsilon_0\frac{d\Phi_E}{dt} = \frac{\mu_0}{2\pi r} i_d.$$

This equation stresses that **B** in the cavity is associated with a displacement current; compare Eq. 34–4. Applying the right-hand rule in Fig. 38–10*a* shows that the displacement current i_d must be directed into the plane of the figure if it is to be associated with the clockwise lines of **B** that are present.

The displacement current is represented in Fig. 38–11*b* by the arrows that point to the right and in Fig. 38–11*a* by the crosses that represent arrows entering the page. Study of Fig. 38–11 shows that the current is continuous, being directed up the



(a) (b)

Fig. 38–11 The cavity of Fig. 38–10 showing (a) the conduction current coming up the walls and displacement current going down the cavity volume and (b) the displacement current (black arrowheads) in the volume of the cavity and the conduction currents (white arrowheads) at the walls. The arrows in each case represent *current densities*. Note the continuity of current (conduction + displacement), that is, it is possible to form closed current loops. (For a *truly* resistanceless cavity the conduction current would lie *entirely* on the surface, requiring a modification of our definition of current density; we leave this small complication for the student to consider.)

walls as a conduction current and then back down through the volume of the cavity as a displacement current. If we apply Ampère's law as extended by Maxwell,

$$\oint \mathbf{B}\cdot d\mathbf{l} = \mu_0(i_d + i), \tag{38–25}$$

to the circular path of radius r_1 in Fig. 38–11*a*, we see that **B** at that path is due entirely to the displacement current, the conduction current i *within the path* being zero.

For the path of radius r_2, the *net* current enclosed is zero because the conduction current in the walls is exactly equal and opposite to the displacement current in the cavity volume. Since i equals i_d in magnitude, but is oppositely directed, it follows from Eq. 38–25 that B must be zero for all points outside the cavity, in agreement with observation.

QUESTIONS

1. Why doesn't the *LC* circuit of Fig. 38–1 simply stop oscillating when the capacitor has been completely discharged?
2. How might you start an *LC* circuit into oscillation with its initial condition being represented by Fig. 38–1*c*? Devise a switching scheme to bring this about.
3. In an oscillating *LC* circuit, assumed resistanceless, what determines (*a*) the frequency and (*b*) the amplitude of the oscillations?
4. Tabulate as many mechanical or electric systems as you can think of that possess a natural frequency, along with the formula for that frequency if given in the text.
5. Resonance in *LC* circuits, judged by Eq. 38–14 and Fig. 38–6, occurs when the frequency ω'' of the "driving force" is exactly equal to the (undamped) natural frequency of the *LC* circuit. In Section 15–10 we saw that resonance for mass-spring systems, judged by Eq. 15–41 and Fig. 15–19, occurs when ω'' is close to, *but not exactly equal to*, the (undamped) natural frequency of the mass-spring system. Is there a failure of the principle of correspondence here?

6. Suppose that, in a low frequency oscillating *LCR* circuit, one could display on the screen of a cathode-ray oscilloscope a potential difference proportional to the temperature of the (specially designed) resistor as it changed with time in response to the heating effect of the oscillating current. What would the connection be between the frequency of this temperature-time waveform and that of the oscillations of current, as measured by the potential difference across the resistor?

7. Discuss the periodic flow of energy, if any, from point to point in an acoustic resonant cavity.

8. Can a given circuit element (a capacitor, say) behave like a "lumped" element under some circumstances and like a "distributed" element under others?

9. List as many (*a*) lumped and (*b*) distributed mechanical oscillating systems as you can.

10. An air-filled acoustic resonant cavity and an electromagnetic resonant cavity of the same size have resonant frequencies that are in the ratio of 10^6 or so. Which has the higher frequency and why?

11. What constructional difficulties would you encounter if you tried to build an *LC* circuit of the type shown in Fig. 38–1 to oscillate (*a*) at 0.01 cps, or (*b*) at 10^{10} cycles/sec?

12. Could an electromagnetic cavity of the type shown in Fig. 38–8 be designed to operate at 60 cycles/sec? If so, give some details of its construction.

13. The "sharpness of tuning" of a copper electromagnetic cavity can be considerably increased by immersing it in liquid air. Explain. Be guided by Fig. 38–6, which shows "tuning curves" for an *LC* circuit.

14. Electromagnetic cavities are often silver plated on the inside. Explain.

15. Why is Faraday's law of induction more familiar than its symmetrical counterpart, Eq. 38–20?

16. Why is the quantity $\epsilon_0 \, d\Phi_E/dt$ referred to as a (displacement) *current?*

17. In Fig. 38–1*c* a displacement current is needed to maintain continuity of current in the capacitor. How can one exist, since there is no charge on the capacitor?

18. Discuss the symmetries that appear between (*a*) the first two and (*b*) the second two of Maxwell's equations.

19. At what parts of the cycle will (*a*) the conduction current and (*b*) the displacement current in the cavity of Fig. 38–8 be zero?

20. Discuss the time variation during one complete cycle of the charges that appear at various points on the inner walls of the oscillating electromagnetic cavity of Fig. 38–8.

PROBLEMS

1. You are given a 10-mh inductor and two capacitors, of 5.0- and 2.0-μf capacitance. What resonant frequencies can be obtained by connecting these elements in various ways?

2. Given a 1.0-mh inductor, how would you make it oscillate at 1.0×10^6 cycles/sec?

3. An inductor is connected across a capacitor whose capacitance can be varied by turning a knob. We wish to make the frequency of the *LC* oscillations vary linearly with the angle of rotation of the knob, going from 2×10^5 cycles/sec to 4×10^5 cycles/sec as the knob turns through 180°. If $L = 1.0$ mh, plot C as a function of angle for the 180° rotation.

4. A 10-henry coil has a resistance of 180 ohms. What size of capacitor must be put in series with it if the combination is to "resonate" when connected to a 60-cycle/sec outlet?

5. Derive the differential equation for an *LC* circuit (Eq. 38–5) using the loop theorem.

6. Derive Eq. 38–12, the differential equation for forced oscillations in an *LCR* circuit, from the principle of conservation of energy.

7. Show that a damped *LC* circuit (see Example 3) loses half its energy to Joule heat in a time given approximately by $0.69\tau_L$, in which τ_L is the inductive time constant.

8. A circuit has $L = 10$ mh and $C = 1.0\ \mu$f. How much resistance must be inserted in the circuit to reduce the (undamped) resonant frequency by 0.01%?

9. Suppose that in an oscillating LCR circuit the amplitude of the charge oscillations drops to one-half its initial value after n cycles. Show that the fractional reduction in the frequency of resonance, caused by the presence of the resistor, is given to a close approximation by

$$\frac{\omega - \omega'}{\omega} = \frac{0.0061}{n^2},$$

which is independent of L, C, or R. Apply to the decay curve of Fig. 38–3.

10. Show that, for low damping, the current in a damped LC circuit is given approximately by

$$i = -q_m\omega' e^{-Rt/2L} \sin(\omega' t + \phi),$$

in which

$$\phi = \tan^{-1} \frac{R}{2L\omega'}.$$

Start from Eq. 38–10.

11. *"Q" for a circuit.* In the damped LC circuit of Example 3 show that the fraction of the energy lost per cycle of oscillation, $\Delta U/U$, is given to a close approximation by $2\pi R/\omega L$. The quantity $\omega L/R$ is often called the "Q" of the circuit (for "quality"). A "high-Q" circuit has low resistance, low fractional energy loss per cycle ($= 2\pi/Q$) and (see Fig. 38–6) a sharp resonance or "tuning" curve.

12. Show that the amplitude of the charge oscillations in an oscillating LCR circuit is given by

$$q_m = \frac{\mathcal{E}_m}{\sqrt{\left(\omega''^2 L - \dfrac{1}{C}\right)^2 + (\omega'' R)^2}}$$

For what value of ω'' will q_m be a maximum?

13. An LCR circuit has $L = 1.0$ henry, $C = 20$ μf, and $R = 20$ ohms. For what frequency ω'' of an applied emf will it resonate with maximum response? At what frequencies will the response be one-half its maximum value?

14. Show that the fractional half-width of the resonance curves of Fig. 38–6 is given, to a close approximation, by

$$\frac{\Delta\omega}{\omega} = \frac{\sqrt{3}R}{\omega L},$$

in which ω is the resonant frequency and $\Delta\omega$ is the width of the resonance peak at $i = \frac{1}{2} i_m$. Note (see Problem 11) that this expression may be written as $\sqrt{3}/Q$ which shows clearly that a "high-Q" circuit has a sharp resonance peak, that is, a small $\Delta\omega/\omega$.

15. A resistor-inductor-capacitor combination R_1, L_1, C_1 (connected in series) exhibits resonance at the same frequency as a second combination R_2, L_2, C_2. If the two combinations are now connected in series, at what frequency would the whole circuit resonate?

16. In Example 5 show that the *displacement current density* j_d is given, for $r < R$, by

$$j_d = \epsilon_0 \frac{dE}{dt}.$$

17. Prove that the displacement current in a parallel-plate capacitor can be written as

$$i_d = C \frac{dV}{dt}.$$

18. You are given a 1.0-μf capacitor. How would you establish an (instantaneous) displacement current of 1.0 amp in the space between its plates?

19. In Example 5 how does the displacement current through a concentric circular loop of radius r vary with r? Consider both $r < R$ and $r > R$.

20. In microscopic terms the principle of continuity of current may be expressed as

$$\oint (\mathbf{j} + \mathbf{j}_d)\cdot d\mathbf{S} = 0,$$

in which $\mathbf{j}$ is the conduction current density and $\mathbf{j}_d$ is the displacement current density. The integral is to be taken over any closed surface; the equation essentially says that whatever current flows into the enclosed volume must also flow out. (*a*) Apply this equation to the surface shown by the dashed lines in Fig. 38–12 shortly after switch S is closed. (*b*) Apply it to various surfaces that may be drawn in the cavity of Fig. 38–11, including some that cut the cavity walls.

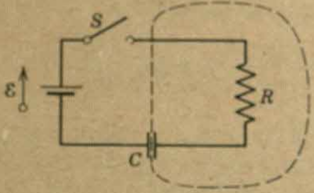


Fig. 38–12

21. A cylindrical electromagnetic cavity 5.0 cm in diameter and 7.0 cm long is oscillating in the mode shown in Fig. 38–8. (*a*) Assume that, for points on the axis of the cavity, $E_m = 10^4$ volts/meter. For such axial points what is the maximum rate $(dE/dt)_m$ at which E changes? (*b*) Assume that the average value of $(dE/dt)_m$, for all points over a cross section of a cavity, is about one-half the value found above for axial points. On this assumption, what is the maximum value of B at the cylindrical surface of the cavity?

22. Collect and tabulate expressions for the following four quantities, considering both $r < R$ and $r > R$. Copy down the derivations side by side and study them as interesting applications of Maxwell's equations to problems having cylindrical symmetry.

(*a*) $B(r)$ for a current i in a long wire of radius R (see Section 34–2).

(*b*) $E(r)$ for a long uniform cylinder of charge of radius R (see Section 28–6; also Problem 14, Chapter 28).

(*c*) $B(r)$ for a parallel-plate capacitor, with circular plates of radius R, in which E is changing at a constant rate (see Section 38–7).

(*d*) $E(r)$ for a cylindrical region of radius R in which a uniform magnetic field B is changing at a constant rate (see Section 35–5).

Electromagnetic Waves *

CHAPTER 39

39-1 Transmission Line

In Chapter 38 we studied electromagnetic energy confined, as a standing wave, to a restricted region of space, the interior of an electromagnetic resonant cavity. Such energy can also be transferred from place to place as a traveling wave. An arrangement of conductors for facilitating such transfers is called a *transmission line.* Figure 39-1 shows one type of line, a *coaxial cable,* its input end being connected to a switch S. For the time being we assume that the cable is infinitely long and that the cable elements have zero resistance.

When S is closed on b, the central and the outer conductors are at the same potential. If the switch is then thrown to a, a potential difference V suddenly appears between these elements. This potential difference does not appear instantaneously all along the line but is propagated with a finite

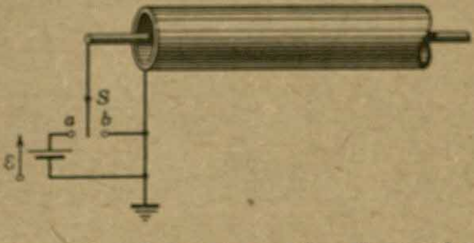


Fig. 39-1 An electromagnetic pulse can be sent along the coaxial cable by throwing switch S from b to a.

---

* This subject is treated in greater depth in Supplementary Topic V at the end of the book.

speed c that will turn out to be exactly that of light, assuming a resistanceless line. Figure 39–2a shows that the potential difference between the conductors at a distance l along the line suddenly rises, at a time given by $t = l/c$, from zero to a value determined by the battery emf. We can also consider the variation of V with position x along the line at a given time t_1 after closing the switch. Figure 39–2b shows such an instantaneous "snapshot." It, too, suggests a traveling "wavefront" moving along the line at speed c. At $t = t_1$ the signal has not yet reached points where $x > ct_1$.

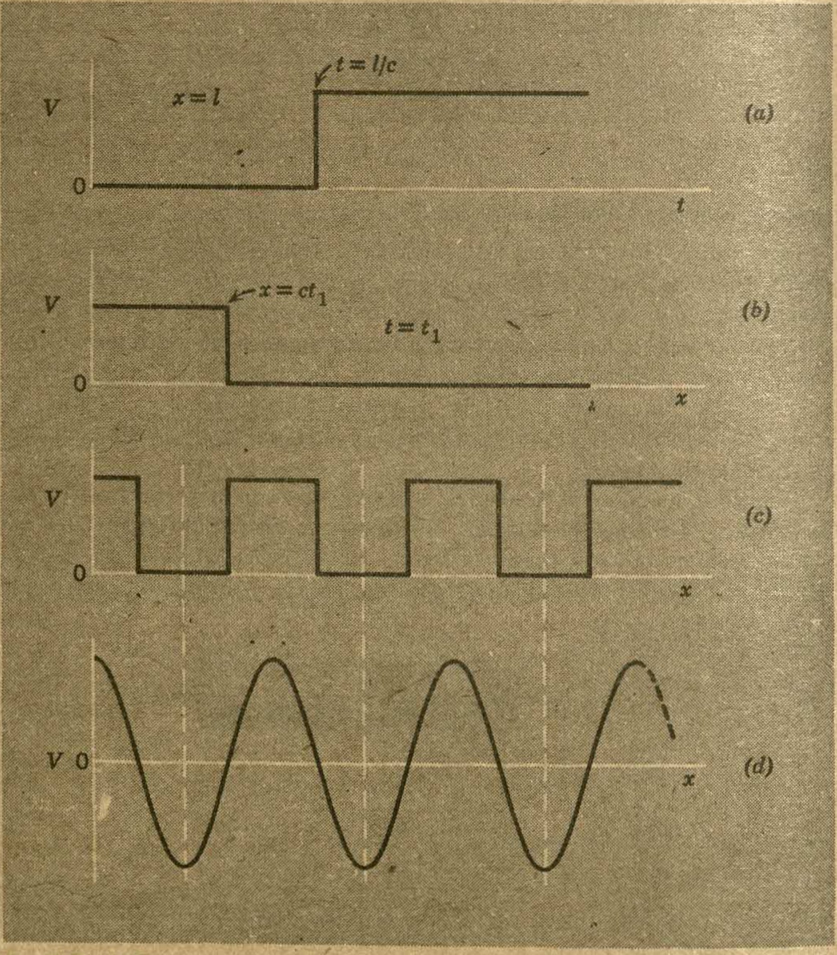


Fig. 39–2 (a) The variation with time of the potential difference between the conductors of a coaxial cable at a distance l from the input end. (b) An instantaneous "snapshot" of the pulse in the cable at a certain time t_1. (c) The waveform if switch S in Fig. 39–1 is periodically thrown between a and b. (d) The waveform if switch S is replaced by an electromagnetic oscillator with a sinusoidal output.

If switch S is periodically thrown from b to a and back again, a wave disturbance like that of Fig. 39-2c is propagated. This suggests that if the battery and switch arrangement is replaced by an electromagnetic oscillator with a sinusoidal output of frequency f a wave like that of Fig. 39-2d will be propagated.

A traveling wave in a resistanceless transmission line will exhibit a wavelength λ given by

$$\lambda = \frac{c}{\nu}.$$

If the oscillator frequency is 60 cycles/sec, the common commercial power frequency, the wavelength is 5×10^6 meters, which is about 3000 miles. At this low frequency traveling waves are not apparent in any line of normal length. By the time the polarity of the oscillator has changed appreciably, the energy fed into the line at the oscillator end has been delivered to the load.

Frequencies in the radio or the microwave range are much higher and the wavelengths correspondingly smaller. Commercial television frequencies as established by the Federal Communications Commission range from 54×10^6 to 980×10^6 cycles/sec. In terms of wavelength this is a range of 5.6 to 0.31 meter. At these wavelengths the patterns of potential difference in the transmission lines used to send television signals across the country can be described aptly as traveling waves. Microwaves, used in radar systems and for communication purposes, have even smaller wavelengths, in the range of about 20 cm to about 0.5 mm.

These considerations suggest another way of viewing the difference between lumped and distributed circuit elements. A system is "distributed" if the wavelength is about the same size as, or less than, the dimensions of the system. If the wavelength is much larger than the dimensions of the system, we are dealing with lumped components. A transmission line 50 meters long would be a lumped system for electromagnetic radiation at 60 cycles/sec ($\lambda = 5 \times 10^6$ meters) but a distributed system at 10^8 cycles/sec ($\lambda = 3$ meters). In a lumped system the circuit analysis is normally carried out in terms of lumped system parameters such as L, C, and R; in a distributed system the analysis is often carried out in terms of the fields that are set up and the charges and currents that are related to them.

▶ **Example 1.** A potential difference given by

$$V_0 = V_m \sin \omega t$$

is applied between the terminals of a long resistanceless transmission line; the frequency ν $(= \omega/2\pi)$ is 3×10^9 cycles/sec. Write an equation for $V(t)$ at a point P which is 1.5 wavelengths down the line from the oscillator.

The general equation for a wave traveling in the x direction (see Eq. 19-10) is

$$V = V_m \sin (\omega t - kx),$$

where $k\ (= 2\pi/\lambda)$ is the wave number. At $x = 0$ this gives correctly the time variation of the input terminal potential difference. At $x = 1.5\lambda$ we have

$$V_P = V_m \sin\left[\omega t - \left(\frac{2\pi}{\lambda}\right)(1.5\lambda)\right] = V_m \sin(\omega t - 3\pi)$$

$$= -V_m \sin \omega t.$$

Thus V_P is always equal in magnitude to V_0 but is opposite in sign. What is the wavelength in this example? ◀

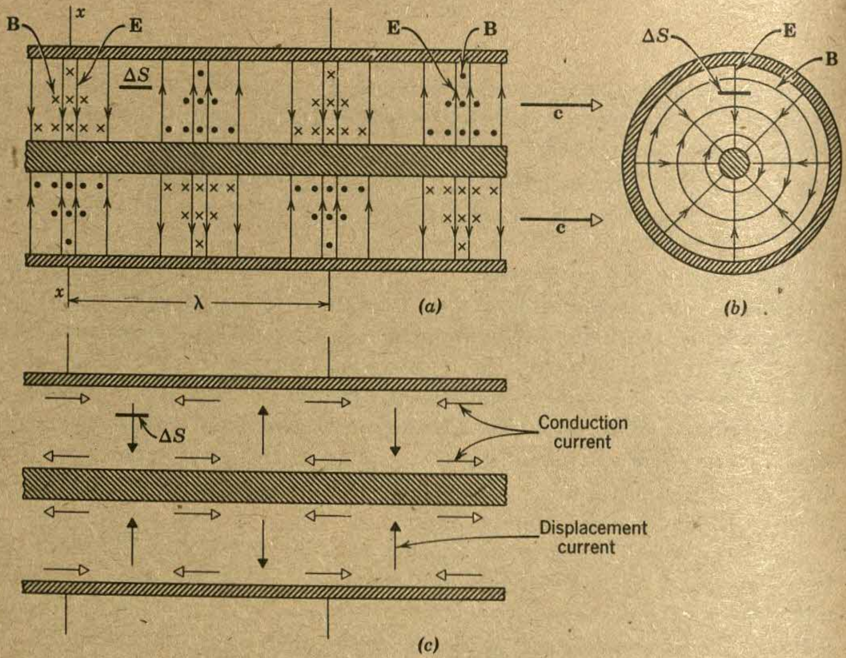


Fig. 39–3 (*a*) The electric and magnetic fields in a coaxial cable, showing a wave traveling to the right at speed *c*. (*b*) A cross-sectional view at a plane through *xx* in (*a*); the wave is emerging from the page. (*c*) Conduction current (open arrows) and displacement currents (filled arrows) associated with the wave in (*a*); the arrows in each case represent current density vectors.

39–2 Coaxial Cable—Fields and Currents

Figures 39–3*a* and *b* are "snapshots" of the electric and magnetic fiel configurations in a coaxial cable. The electric field is radial and the magneti field forms concentric lines about the central conductor. The entire patter moves along the line, assumed resistanceless, at speed *c*.

The field patterns in this figure obey the *boundary condition* required for line that is assumed to be resistanceless, namely, that **E** for all points either conducting surface has no tangential component (see p. 957). Tl field patterns can be deduced mathematically from Maxwell's equations imposing this requirement. The configuration shown is the simplest of many different wave patterns that can travel along the line. The coaxial cable, unlike the electromagnetic cavity of Fig. 38–8, is not a resonant device. The angular frequency ω of waves that travel along it can be varied continuously, as is the case for all traveling waves, such as transverse waves in a long stretched cord.

Figure 39–3*c* shows the currents in the cable at the instant corresponding to Figs. 39–3*a* and *b*. The arrows parallel to the cable axis represent conduction currents in the central and the outer conductors. The vertical arrows with filled heads represent displacement currents that exist in the space between the conductors. Note that the conduction current and the displacement current arrows form closed loops, preserving the concept of the continuity of current.

▶ **Example 2.** Verify that the displacement current represented in Fig. 39–3*c* is consistent with the pattern of **B** and **E** shown in Fig. 39–3*a*.

Consider a small surface element $\Delta\mathbf{S}$ shown edge-on in Fig. 39–3; it is shown as viewed from above in Fig. 39–4. This hypothetical element is stationary with respect to the cable while the field configuration moves through it at speed c. Figure 39–4*a* shows the electric lines of force in the vicinity of this element. It is clear from symmetry that, at the instant shown, the net flux Φ_E through this area is zero. However, even though Φ_E is zero in magnitude, it is, at this instant, *changing at its most rapid*

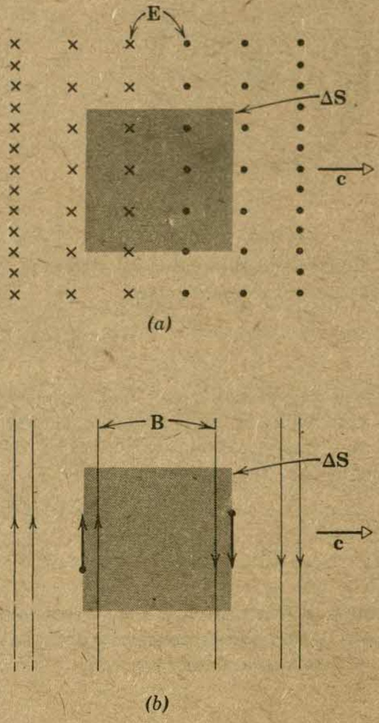


Fig. 39–4 The area element S in Fig. 39–3 enlarged and viewed from above, showing the adjacent (*a*) electric and (*b*) magnetic fields.

rate, since $\mathbf{E}$ at the element ΔS is at the very moment of reversing its direction as the wave moves through. Thus the displacement current, which is given by

$$i_d = \epsilon_0 \frac{d\Phi_E}{dt},$$

also has its maximum value.

Figure 39–4*b* shows the magnetic field in the vicinity of the element of area. Let us apply the generalized form of Ampère's law,

$$\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0(i + i_d),$$

to the element. The conduction current i is zero since no charge is transported through ΔS. The displacement current i_d is not zero, having, in fact, its maximum value. Thus, since the right side of this equation $(= \mu_0 i_d)$ does not vanish, the left side must not vanish. Study of Fig. 39–4*b* shows that $\oint \mathbf{B} \cdot d\mathbf{l}$ around the boundary of this square has, indeed, a nonzero value. Thus the field and displacement current configurations of Figs. 39–3*a*–*c* are consistent. We have not discussed the direction of i_d; that is, does it point into the plane of Fig. 39–4, as Fig. 39–3*c* asserts, or out of it? We leave this as a question for the student. He may be guided by considering the direction of the displacement current in Figs. 38–9*a* and *b*.

Example 3. Show that the conduction currents in Fig. 39–3*c* are appropriately related to the magnetic field pattern.

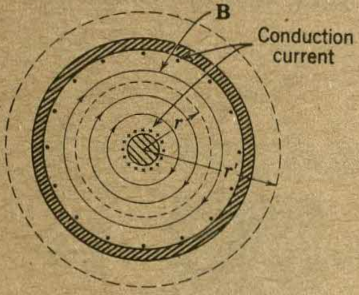


Fig. 39–5 Example 3. The coaxial cable of Fig. 39–3*b*, showing the conduction currents in the central and outer conductors. The wave is emerging from the page.

Figure 39–5 shows a cross section of the cable at a plane through xx in Fig. 39–3*a*. Let us apply Ampère's law,

$$\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0(i + i_d),$$

to the ring of radius r. The displacement current through this ring is zero, this current being at right angles to the central conductor. The conduction current i in the central conductor, shown by the $\times$'s in the figure, does pass through this ring so that the equation becomes

$$(B)(2\pi r) = \mu_0 i,$$

or

$$B = \frac{\mu_0 i}{2\pi r}.$$

Note that $\mathbf{B}$ is related to the current in the central conductor by the usual right-hand rule and that the expression for B is that found earlier (Eq. 34–4) for a long straight wire carrying a steady current.

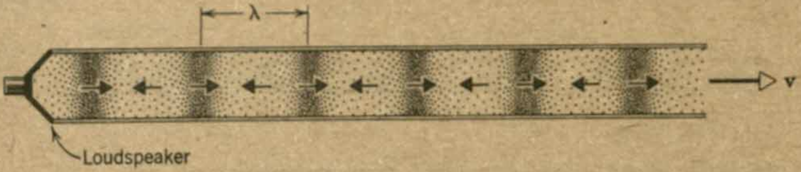


Fig. 39-6 An acoustic transmission line, showing a sound wave traveling to the right. The small arrows with filled heads show the directed drift velocities for small volume elements of the gas. Compare Fig. 39-3*a*.

If we apply Ampère's law to the large ring of radius r', we must put the *net* conduction current equal to zero because the current in the outer conductor, shown by the dots, is equal and opposite to that in the inner conductor. This means that **B** must be zero for points outside the cable, in agreement with experiment. ◀

It is interesting to compare the electromagnetic oscillations in a typical *traveling* wave, such as that of Fig. 39-3, to those in a cavity resonator, such as that of Fig. 38-8. The latter oscillations are an electromagnetic *standing* wave. In a traveling wave **E** and **B** are in phase, which means that at a given position along the transmission line they reach their maxima at the same time. However, Fig. 38-8 shows that at a given position in a cavity resonator **E** and **B** reach their maxima one-fourth of a cycle apart; they are 90° out of phase.

A complete analogy exists in mechanical systems. In the acoustic resonator of Fig. 38-7 the time variations of pressure and velocity for the standing acoustic wave are also 90° out of phase, in exact correspondence to the electromagnetic cavity oscillations of Fig. 38-8. The acoustic analogy to a transmission line (Fig. 39-6) would be an infinitely long gas-filled tube, one end being connected to an acoustic oscillator such as a loud-speaker. The entire configuration of Fig. 39-6 moves to the right with speed v. The pressure variations, suggested by the dots, and the instantaneous velocities, suggested by the arrows, are *in phase*, just as are **E** and **B** in the coaxial cable of Fig. 39-3.

39-3 Waveguide

It is possible to send electromagnetic waves through a hollow metal pipe. We assume that the inner walls of such a pipe, or *waveguide* as it is called, are resistanceless and that the cross section is rectangular.

Figure 39-7 shows a typical electric and magnetic field pattern. We imagine that a microwave oscillator is connected to the left end and sends electromagnetic energy down the guide. Figure 39-7*a* shows a side view of the guide and Fig. 39-7*b*, a top view; Fig. 39-7*c* shows the cross section. As for the coaxial cable, the field patterns are such that **E** has no tangential component for any point on the inner surface of the guide. The fields **E** and **B** are in phase, again like the coaxial cable.

As for all traveling waves, the angular frequency ω of electromagnetic waves traveling down a guide can be varied continuously. In a waveguide of given dimensions, however, there exists, for every mode of transmission, that is, for every pattern of **E** and **B**, a so-called *cutoff frequency* ω_0. A given guide will not transmit waves in a given mode if their frequency is below the cutoff value for that mode in that guide. The field patterns of Fig. 39-7 show the *dominant mode* for a rectangular guide; this the mode with the lowest cutoff frequency. Given the frequency ω of electromagnetic waves to be transmitted, it is common practice to select a guide whose dimensions are such that ω is larger than the cutoff frequency ω_0 for the dominant mode

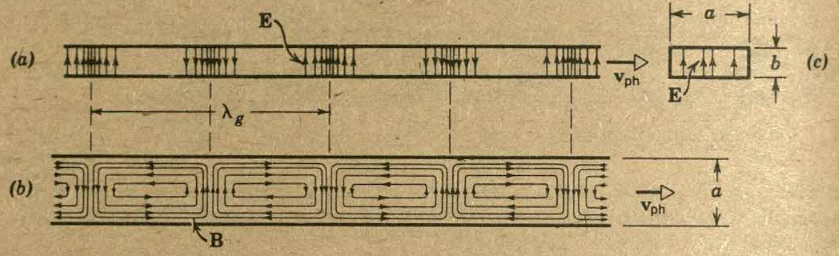


Fig. 39–7 A waveguide, showing (*a*) a side view of the lines of **E**, (*b*) a top view of the lines of **B**, and (*c*) a cross-sectional view of the lines of **E**. In (*c*) the wave is emerging from the page. For simplicity the lines of **B** are not shown in (*a*) and (*c*), nor are the lines of **E** shown in (*b*).

but smaller than the cutoff frequencies of all other modes. Under these conditions the dominant mode of propagation is the only one possible.

In a (resistanceless) coaxial cable the wave patterns travel at speed c. In the acoustic transmission line of Fig. 39–6 (assumed "resistanceless") the waves also travel at a speed v, which is the same as the propagation speed in an infinite medium. In a waveguide, however, the speed is *not* c. In waveguides we must distinguish between (*a*) the *phase speed* v_{ph}, which is the speed at which the wave patterns of Fig. 39–7 travel, and (*b*) the *group speed* v_{gr}, which is the speed at which electromagnetic energy or information-carrying "signals" travel along the guide. These speeds, which are identical for electromagnetic waves in a coaxial cable and for acoustic waves in a tube, are different for waves in a waveguide.

The phase speed is not directly measurable. The wave pattern is a repetitive structure, and there is no way to distinguish one wave maximum from another. The waves can be observed to enter one end of the guide and to leave at the other, but there is no way to identify a particular wave maximum so that its passage down the guide can be timed. We can put a "signal" on the wave by increasing the power level of the oscillator for a short time. This power pulse could be timed as it passes through the guide, but there is no guarantee that it travels at the same speed as the wave pattern and, indeed, it does not. The speed of such signals or markers is the speed at which *energy* is propagated, that is, the group speed.

From Maxwell's equations it can be shown that the phase speed and the group speed for the mode of Fig. 39–7 are

$$v_{\mathrm{ph}} = \frac{c}{\sqrt{1 - \left(\frac{\lambda}{2a}\right)^2}} \tag{39–1}$$

and

$$v_{\mathrm{gr}} = c\sqrt{1 - \left(\frac{\lambda}{2a}\right)^2}, \tag{39–2}$$

in which a is the width of the guide and λ the free-space wavelength. Note that as $a \to \infty$, which corresponds to free-space conditions, $v_{\mathrm{ph}} = v_{\mathrm{gr}} = c$.

The phase speed v_{ph} is *greater* than the velocity of light, the group speed v_{gr} being correspondingly less. In relativity theory we learn that no speed at which signals or energy travel can be faster than that of light. However, signals or energy cannot be transmitted down a guide at speeds exceeding c; they travel with speed v_{gr} which is always less than c so there is no conflict with the theory of relativity.

The wavelength λ in Eqs. 39-1 and 39-2 is the wavelength that would be measured for the oscillations in free space, that is,

$$\lambda = \frac{c}{\nu}, \tag{39-3}$$

where c is the speed in free space and ν is the frequency. For waves of a given frequency, the wavelength exhibited in a guide (λ_g) must differ from the free-space wavelength λ because the speed v_{ph} has changed. The so-called *guide wavelength* λ_g is given by

$$\lambda_g = \frac{v_{\text{ph}}}{\nu} = \frac{v_{\text{ph}}}{c/\lambda} = \lambda\frac{v_{\text{ph}}}{c}.$$

From Eq. 39-1 this yields

$$\lambda_g = \frac{\lambda}{\sqrt{1 - \left(\frac{\lambda}{2a}\right)^2}}. \tag{39-4}$$

Thus the guide wavelength, which is the wavelength exhibited by the field patterns in Fig. 39-7, is larger than the free-space wavelength.

▶ **Example 4.** What must be the width a of a rectangular guide such that the energy of electromagnetic radiation whose free-space wavelength is 3.0 cm travels down the guide (*a*) at 95% of the speed of light? (*b*) At 50% of the speed of light?

From Eq. 39-2 we have

$$v_{\text{gr}} = 0.95c = c\sqrt{1 - \left(\frac{\lambda}{2a}\right)^2}.$$

Solving for a yields $a = 4.8$ cm; repeating for $v_{\text{gr}} = 0.50c$ yields $a = 1.7$ cm.

This formula illustrates the cutoff phenomenon described above. If $\lambda = 2a$, then $v_{\text{gr}} = 0$ and energy cannot travel down the guide. For the radiation considered in this example $\lambda = 3.0$ cm, so that the guide must have a width a of *at least* $\frac{1}{2}$ x 3.0 cm = 1.5 cm if it is to transmit this wave. The guide whose width we calculated in (*a*) above can transmit radiations whose free-space wavelength is 2 x 4.8 cm = 9.6 cm *or less.* ◀

39-4 Radiation

The acoustic transmission line of Fig. 39-6 cannot be infinitely long. Its far end may be sealed by a solid cap or left open, or it may have a flange, a horn, or some similar device mounted on it. If the far end is not sealed, energy will escape into the medium beyond. This is called *acoustic radiation.* In general, some energy will also be reflected back down the transmission line. If acoustic radiation is desirable, the designer's task is to fashion a termination (that is, an "acoustic antenna") for the transmission line such that the smallest possible fraction of the incident energy will be reflected back down the line. Such a termination might take the form of a flared horn. Acoustic radiation, of course, requires a medium such as air in order to be propagated.

An electromagnetic transmission line such as a coaxial cable or a waveguide can also be terminated in many ways, and energy can escape from the end of the line into the space beyond. In contrast to sound waves, a physical

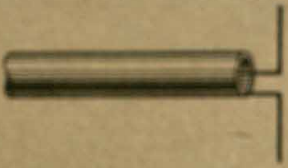


Fig. 39-8 An electric dipole antenna on the end of a coaxial cable.

medium is not required. Thus electromagnetic energy can be radiated from the end of the transmission line, to form a traveling electromagnetic wave in free space.

Figure 39-8 shows an effective termination for a coaxial cable; it consists of two wires arranged as shown and is called an *electric dipole antenna*. The potential difference between the two conductors alternates sinusoidally as the wave reaches them, the effect being that of an electric dipole whose dipole moment **p** varies with time.

Figure 39-9 shows such a dipole, represented by two equal and opposite charges. Its dipole moment, represented by the arrows marked **p** in the

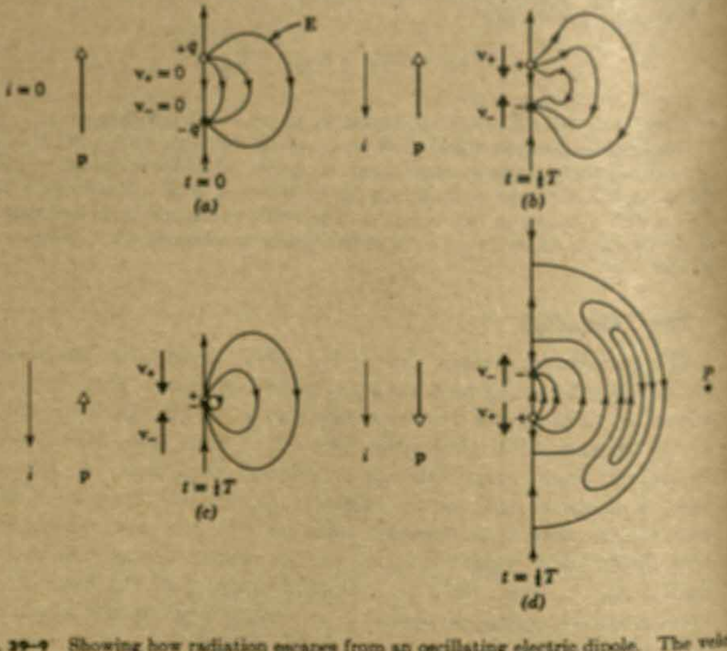


Fig. 39-9 Showing how radiation escapes from an oscillating electric dipole. The velocities of the charges, the electric dipole moment, the equivalent current, and the lines of E are shown at four stages of oscillation, one-eighth of a period apart. The lines of B are not shown. (Adapted from Attwood, *Electric and Magnetic Fields*, John Wiley and Sons, third edition, 1949.)

figure, oscillates sinusoidally as the charges oscillate; in Fig. 39-9c it is on the verge of reversing its direction. The four views, one-eighth of a cycle apart, show how the electric lines of force break away from the dipole and form closed loops that travel through free space with speed c. The oscillating charges constitute a current, which is represented by the arrows marked i in Fig. 39-9. These oscillating currents generate a field of **B** which, for simplicity, is not represented in the figure. The lines of **B** as well as those of **E** also form closed loops that move away from the dipole with speed c. These traveling electric and magnetic fields, which, as we shall see, are strongly interdependent, constitute *electromagnetic radiation*.

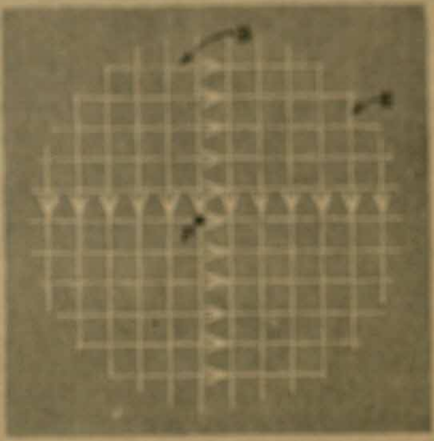


Fig. 39-10 An instantaneous view of an electromagnetic wave as seen by an observer at P in Fig. 39-9d.

Figure 39-10 shows a part of the wavefront as it would appear to an observer at point P in Fig. 39-9d. The wave is moving directly out of the page. One-half a period later the observer at P will see a field pattern like that of Fig. 39-10, except that the directions of both the electric and the magnetic fields will be reversed. The speed c of the wave in free space is given by $c = \nu\lambda$, which can be written as

$$c = \frac{\omega}{k}, \tag{39-5}$$

where ω, the angular frequency, and k, the wave number, are related to the frequency ν and the wavelength λ by

$$\omega = 2\pi\nu \quad \text{and} \quad k = 2\pi/\lambda.$$

39-5 Traveling Waves and Maxwell's Equations

In earlier sections we have postulated the existence of certain magnetic and electric field distributions, in resonant cavities, coaxial cables, and waveguides, and we have shown that these postulated distributions are consistent with Maxwell's equations, as are the distributions of conduction and displacement currents associated with the fields. The student who pursues his studies of electromagnetism will learn how to derive mathematical expressions for **E** and **B** by subjecting Maxwell's equations to the boundary conditions appropriate to the problem at hand. In this section we continue our program by showing that the postulated patterns of **E** and **B** for a traveling electromagnetic wave are completely consistent with Maxwell's equations. In doing so, we will be able to show that the speed of such waves in free space is that of visible light and thus that visible light is itself an electromagnetic wave.

If the observer at P in Fig. 39-9*d* is a considerable distance from the source, the *wavefronts* described by the electric and magnetic fields that reach him (see Fig. 39-10) will be planes and the wave that moves past him will be a *plane wave* (see Section 19-2). Figure 39-11 shows a "snapshot" of a plane wave traveling in the x direction. The lines of **E** are parallel to the z axis and those of **B** are parallel to the y axis. The values of **B** and **E** for this wave depend only on x and t (not on y or z). We postulate that they are given in magnitude by

$$B = B_m \sin (kx - \omega t) \tag{39-6}$$

and

$$E = E_m \sin (kx - \omega t). \tag{39-7}$$

Figure 39-12 shows two sections through the three-dimensional diagram of Fig. 39-11. In Fig. 39-12*a* the plane of the page is the xz plane and in

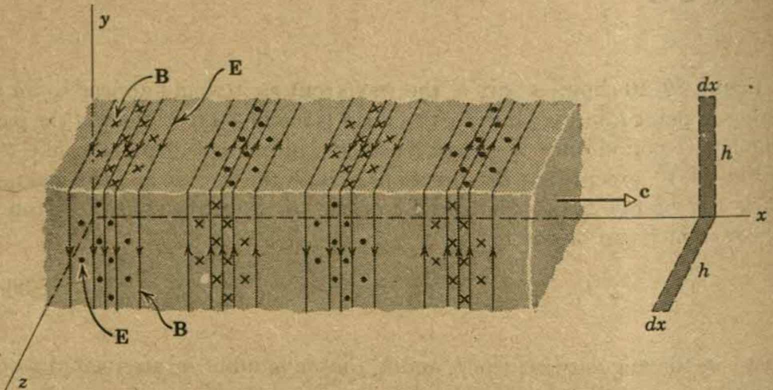


Fig. 39-11 A plane electromagnetic wave traveling to the right at speed c. Lines of **B** are parallel to the y axis; those of **E** are parallel to the z axis. The shaded rectangles on the right refer to Fig. 39-12.

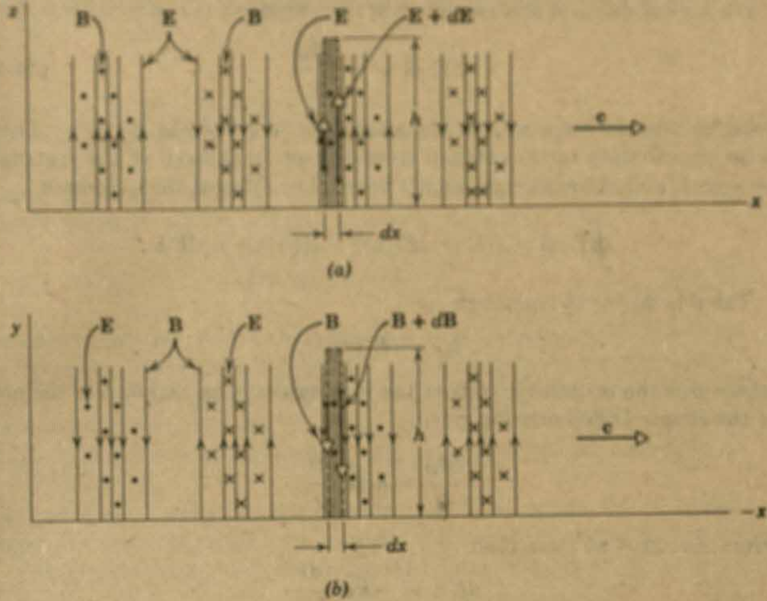


Fig. 39-12 The wave of Fig. 39-11 viewed (*a*) in the *xz* plane and (*b*) in the *xy* plane.

Fig. 39-12*b* it is the *xy* plane. Note that, as for the traveling waves in a coaxial cable (Fig. 39-3*a*) and in a waveguide (Fig. 39-7), **E** and **B** are in phase, that is, at any point through which the wave is moving they reach their maximum values at the same time.

The shaded rectangle of dimensions dx and h in Fig. 39-12*a* is fixed in space. As the wave passes over it, the magnetic flux Φ_B through the rectangle will change, which will give rise to induced electric fields around the rectangle, according to Faraday's law of induction. These induced electric fields are, in fact, simply the electric component of the traveling wave.

Let us apply Lenz's law to this induction process. The flux Φ_B for the shaded rectangle of Fig. 39-12*a* is *decreasing* with time because the wave is moving through the rectangle to the right and a region of weaker magnetic fields is moving into the rectangle. The induced field will act to oppose this change, which means that if we imagine that the boundary of the rectangle is a conducting loop a *counterclockwise* induced current would appear in it. This current would produce a field of **B** that, within the rectangle, would point out of the page, thus opposing the decrease in Φ_B. There is, of course, no conducting loop, but the net induced electric field **E** does indeed act counterclockwise around the rectangle because $E + dE$, the magnitude of **E** at the right edge of the rectangle, is greater than E, the magnitude of **E** at the left edge. Thus the electric field configuration is entirely consistent with the concept that it is induced by the changing magnetic field.

For a more detailed analysis let us apply Faraday's law of induction, or

$$\oint \mathbf{E}\cdot d\mathbf{l} = -\frac{d\Phi_B}{dt}, \tag{39–8}$$

going counterclockwise around the shaded rectangle of Fig. 39–12*a*. There is no contribution to the integral from the top or bottom of the rectangle because **E** and *d***l** are at right angles here. The integral then becomes

$$\oint \mathbf{E}\cdot d\mathbf{l} = [(E + dE)(h)] - [(E)(h)] = dE\, h.$$

The flux Φ_B for the rectangle is

$$\Phi_B = (B)(dx\, h),$$

where B is the magnitude of **B** at the rectangular strip and $dx\, h$ is the area of the strip. Differentiating gives

$$\frac{d\Phi_B}{dt} = h\, dx\, \frac{dB}{dt}.$$

From Eq. 39–8 we have then

$$dE\, h = -h\, dx\, \frac{dB}{dt},$$

or

$$\frac{dE}{dx} = -\frac{dB}{dt}. \tag{39–9}$$

Actually, both B and E are functions of x and t; see Eqs. 39–6 and 39–7. In evaluating dE/dx, it is assumed that t is constant because Fig. 39–12*a* is an "instantaneous snapshot." Also, in evaluating dB/dt it is assumed that x is constant since what is required is the time rate of change of B at a particular place, the strip in Fig. 39–12*a*. The derivatives under these circumstances are called *partial derivatives,* and a somewhat different notation is used for them; see the footnote on p. 725. In this notation Eq. 39–9 becomes

$$\frac{\partial E}{\partial x} = -\frac{\partial B}{\partial t}. \tag{39–10}$$

The minus sign in this equation is appropriate and necessary, for although E is increasing with x at the site of the shaded rectangle in Fig. 39–12*a*, B is decreasing with t. Since $E(x,t)$ and $B(x,t)$ are known (see Eqs. 39–6 and 39–7), Eq. 39–10 reduces to

$$kE_m \cos(kx - \omega t) = \omega B_m \cos(kx - \omega t),$$

or (see Eq. 39–5)

$$\frac{\omega}{k} = \frac{E_m}{B_m} = c. \tag{39–11a}$$

Thus the speed of the wave c is the ratio of the amplitudes of the electric and the magnetic components of the wave. From Eqs. 39–6 and 39–7 we see that the ratio of amplitudes is the same as the ratio of instantaneous values, or

$$E = cB. \tag{39-11b}$$

This important result will be useful in later sections.

We now turn our attention to Fig. 39-12b, in which the flux Φ_E for the shaded rectangle is decreasing with time as the wave moves through it. According to Maxwell's third equation (with $i = 0$, because there are no conduction currents in a traveling electromagnetic wave),

$$\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0 \epsilon_0 \frac{d\Phi_E}{dt}, \tag{39-12}$$

this changing flux will induce a magnetic field at points around the periphery of the rectangle. This induced magnetic field is simply the magnetic component of the electromagnetic wave. Thus, as in the cavity resonator of Section 38-10, the electric and the magnetic components of the wave are intimately connected with each other, each depending on the time rate of change of the other.

Comparison of the shaded rectangles in Fig. 39-12 shows that for each the appropriate flux, Φ_B or Φ_E, is *decreasing* with time. However, if we proceed counterclockwise around the upper and lower shaded rectangles, we see that $\oint \mathbf{E} \cdot d\mathbf{l}$ is *positive*, whereas $\oint \mathbf{B} \cdot d\mathbf{l}$ is *negative*. This is as it should be. If the student will compare Figs. 35-10 and 38-9a, he will be reminded that although the fluxes Φ_B and Φ_E in those figures are changing with time in the same way (both are increasing) the lines of the induced fields, **E** and **B**, respectively, circulate in opposite directions.

The integral in Eq. 39-12, evaluated by proceeding counterclockwise around the shaded rectangle of Fig. 39-12b, is

$$\oint \mathbf{B} \cdot d\mathbf{l} = [-(B + dB)(h)] + [(B)(h)] = -h\,dB,$$

where B is the magnitude of **B** at the left edge of the strip and $B + dB$ is its magnitude at the right edge.

The flux Φ_E through the rectangle of Fig. 39-12b is

$$\Phi_E = (E)(h\,dx).$$

Differentiating gives
$$\frac{d\Phi_E}{dt} = h\,dx \frac{dE}{dt}.$$

Equation 39-12 can thus be written

$$-h\,dB = \mu_0 \epsilon_0 \left(h\,dx \frac{dE}{dt} \right)$$

or, substituting partial derivatives,

$$-\frac{\partial B}{\partial x} = \mu_0 \epsilon_0 \frac{\partial E}{\partial t}. \tag{39-13}$$

Again, the minus sign in this equation is appropriate and necessary, for, although B is increasing with x at the site of the shaded rectangle in Fig. 39–12b, E is decreasing with t.

Combining this equation with Eqs. 39–6 and 39–7 yields

$$-kB_m \cos (kx - \omega t) = -\mu_0 \epsilon_0 \omega E_m \cos (kx - \omega t),$$

or (see Eq. 39–5)

$$\frac{E_m}{B_m} = \frac{k}{\mu_0 \epsilon_0 \omega} = \frac{1}{\mu_0 \epsilon_0 c}. \tag{39–14}$$

Eliminating E_m/B_m between Eqs. 39–11a and 39–14 yields

$$c = \frac{1}{\sqrt{\mu_0 \epsilon_0}}. \tag{39–15}$$

Substituting numerical values yields

$$c = \frac{1}{\sqrt{(4\pi \times 10^{-7} \text{ weber/amp-m})(8.9 \times 10^{-12} \text{ coul}^2/\text{nt-m}^2)}}$$

$$= 3.0 \times 10^8 \text{ meters/sec}, \tag{39–16}$$

which is the speed of light in free space! This emergence of the speed of light from purely electromagnetic considerations is the crowning achievement of Maxwell's electromagnetic theory. Maxwell made this prediction before radio waves were known and before it was realized that light was electromagnetic in nature. His prediction led to the concept of the electromagnetic spectrum, which we discuss in Chapter 40, and to the discovery of radio waves by Heinrich Hertz in 1890. It made it possible to discuss optics as a branch of electromagnetism and to derive its fundamental laws from Maxwell's equations.

A conclusion as fundamental as Eq. 39–15 must be subject to rigorous experimental verification. Of the three quantities in that equation, one, μ_0, has an assigned value, namely, $4\pi \times 10^{-7}$ weber/amp-m. The speed of light c is one of the most precisely measured physical constants, having the presently accepted value of 2.997924×10^8 meters/sec. The remaining quantity, ϵ_0, can be measured by making measurements on an accurately constructed parallel-plate capacitor, as described in Section 30–2. The best measured value, by Rosa and Dorsey of the National Bureau of Standards (U.S.A.) in 1906, is 8.84025×10^{-12} coul^2/nt-m^2. To this accuracy Eq. 39–15 is completely verified. Our confidence in electromagnetic theory, bolstered by numerous successful predictions and agreements with experiment, is now such that we reverse the foregoing procedure and calculate our presently accepted value of ϵ_0 (see Appendix A) from the measured speed of light, using Eq. 39–15.

39–6 The Poynting Vector

One of the important characteristics of an electromagnetic wave is that it can transport energy from point to point. As we show below, the rate of energy flow per unit area in a plane electromagnetic wave can be described by a vector **S**, called the *Poynting vector* after John Henry Poynting (1852–1914), who first pointed out its properties. We define **S** from

$$\mathbf{S} = \frac{1}{\mu_0}\,\mathbf{E}\times\mathbf{B}. \tag{39-17}$$

In the mks system **S** is expressed in watts/meter^2; the direction of **S** gives the direction in which the energy moves. The vectors **E** and **B** refer to their instantaneous values at the point in question. If Eq. 39-17 is applied to the traveling plane electromagnetic wave of Fig. 39-11, it is clear that **E** × **B**, hence **S**, point in the direction of propagation. Note, too, that **S** points parallel to the axis for all points in the coaxial cable of Fig. 39-3.

We get meaningful results if we extend the Poynting vector concept to other electromagnetic situations involving either traveling or standing electromagnetic waves, as we will see in Examples 5 and 6. If we extend it to circuit situations involving steady or almost steady currents and lumped circuit elements, we are led to some interesting conclusions, which we explore in Problems 11, 17, and 18.

▶ **Example 5.** Analyze energy flow in the cavity of Fig. 38-8, using the Poynting vector.

Study of Fig. 39-13 shows that when the energy is all electric (Figs. 39-13*a* and *e*) it is concentrated along the axis, because this is the region in which **E** has its maximum

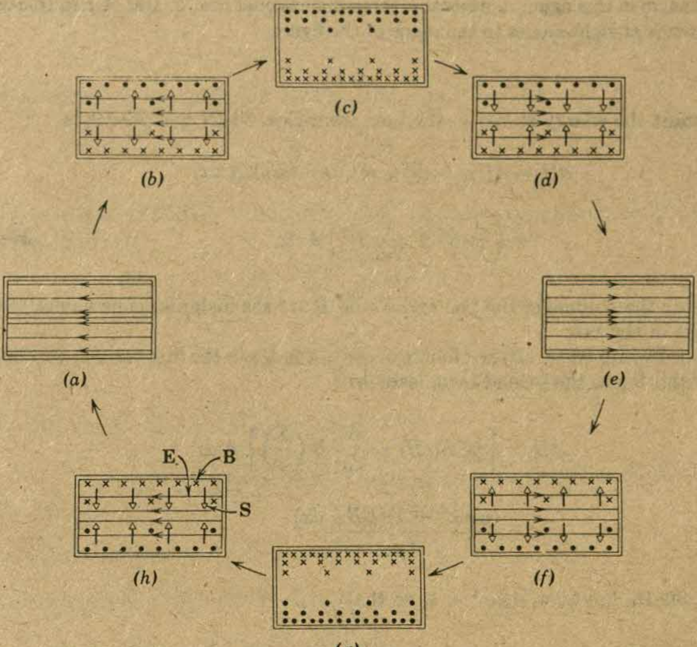


Fig. 39-13 Example 5. Energy surges back and forth periodically between the central region of the cavity and the region near the walls, as indicated by the Poynting vector **S**

value. When the energy is all magnetic (Figs. 39–13*c* and *g*), it is concentrated near the walls. Thus the energy surges back and forth periodically between the central region and the region near the walls. The figure shows, by the open arrows, the direction of **S** at various points in the cavity and at various times in the cycle. Note that **S** equals zero for Figs. 39–13*a*, *c*, *e*, and *g*, which is appropriate because at these instants of time the field configurations are momentarily stationary and energy is not flowing. A pendulum bob at the end of its swing and at the bottom of its trajectory forms a mechanical analogy. The student should verify from Eq. 39–17 that these arrows point in the correct directions. ◀

Figure 39–14 can be used to derive the Poynting relation for the special case of a traveling plane electromagnetic wave. It shows a cross section of a traveling plane wave, along with a thin "box" of thickness *dx* and area *A*. The box, a mathematical construction, is fixed with respect to the axes while the wave moves through it.

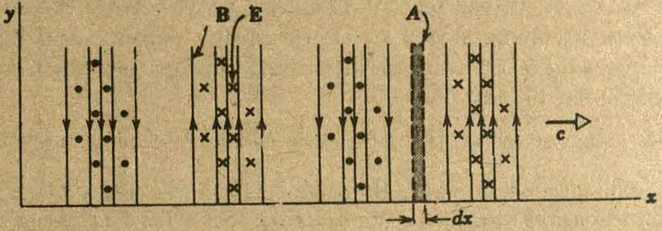


Fig. 39–14 A plane wave is traveling to the right at speed *c*; compare Fig. 39–12*b*. The dashed rectangle in this figure represents a three-dimensional box, of area *A* and thickness *dx*, that extends at right angles to the plane of the figure.

At any instant the energy stored in the box, from Eqs. 30–27 and 36–19, is

$$dU = dU_E + dU_B = (u_E + u_B)(A\,dx)$$

$$= \left(\tfrac{1}{2}\epsilon_0 E^2 + \frac{1}{2\mu_0} B^2\right) A\,dx, \tag{39–18}$$

where $A\,dx$ is the volume of the box and E and B are the instantaneous values of the field vectors in the box.

Using Eq. 39–11*b* ($E = cB$) to eliminate *one* of the E's in the first term in Eq. 39–18 and *one* of the B's in the second term leads to

$$dU = \left[\tfrac{1}{2}\epsilon_0 E(cB) + \frac{1}{2\mu_0} B\left(\frac{E}{c}\right)\right] A\,dx$$

$$= \frac{(\mu_0\epsilon_0 c^2 + 1)(EBA\,dx)}{2\mu_0 c}$$

From Eq. 39–15, however, $\mu_0\epsilon_0 c^2 = 1$, so that

$$dU = \frac{EBA\,dx}{\mu_0 c}.$$

This energy dU will pass through the right face of the box in a time dt equal to dx/c. Thus the energy per unit area per unit time, which is S, is given by

$$S = \frac{dU}{dt\,A} = \frac{EBA\,dx}{\mu_0 c\,(dx/c)\,A} = \frac{1}{\mu_0}EB.$$

This is exactly the prediction of the more general relation Eq. 39-17 for a traveling plane wave.

This relation refers to values of S, E, and B at any instant of time. We are usually more interested in the *average* value of S, taken over one or more cycles of the wave. An observer making intensity measurements on a wave moving past him would measure this average value $\bar{S}$. We can easily show (see Example 6) that $\bar{S}$ is related to the *maximum* values of E and B by

$$\bar{S} = \frac{1}{2\mu_0}E_mB_m.$$

▶ **Example 6.** An observer is at a great distance r from a point light source whose power output is P_0. Calculate the magnitudes of the electric and the magnetic fields. Assume that the source is monochromatic, that it radiates uniformly in all directions, and that at distant points it behaves like the traveling plane wave of Fig. 39-11.

The power that passes through a sphere of radius r is $(\bar{S})(4\pi r^2)$, where $\bar{S}$ is the *average* value of the Poynting vector at the surface of the sphere. This power must equal P_0, or

$$P_0 = \bar{S}4\pi r^2.$$

From the definition of **S** (Eq. 39-17), we have

$$\bar{S} = \overline{\left(\frac{1}{\mu_0}EB\right)}.$$

Using the relation $E = cB$ (Eq. 39-11b) to eliminate B leads to

$$\bar{S} = \frac{1}{\mu_0 c}\overline{E^2}.$$

The average value of E^2 over one cycle is $\frac{1}{2}E_m{}^2$, since E varies sinusoidally (see Eq. 39-7). This leads to

$$P_0 = \left(\frac{E_m{}^2}{2\mu_0 c}\right)(4\pi r^2),$$

or

$$E_m = \frac{1}{r}\sqrt{\frac{P_0\mu_0 c}{2\pi}}.$$

For $P_0 = 10^3$ watts and $r = 1.0$ meter this yields

$$E_m = \frac{1}{(1.0m)}\sqrt{\frac{(10^3\text{ watts})(4\pi \times 10^{-7}\text{ weber/amp-m})(3 \times 10^8\text{ meters/sec})}{2\pi}}$$

$$= 240\text{ volts/meter}.$$

The relationship $E_m = cB_m$ (Eq. 39-11a) leads to

$$B_m = \frac{E_m}{c} = \frac{240\text{ volts/meter}}{3 \times 10^8\text{ meters/sec}} = 8 \times 10^{-7}\text{ weber/meter}^2.$$

Note that E_m is appreciable as judged by ordinary laboratory standards but that B_m (= 0.008 gauss) is quite small.

QUESTIONS

1. In the coaxial cable of Fig. 39–1, what are the directions of the conduction current (*a*) in the central conductor and (*b*) in the outer conductor, shortly after the switch is thrown to position *a*? Consider points that have been reached by the wavefront of Fig. 39–2*a* and *b* and those that have not.

2. Compare a coaxial cable and a waveguide, used as a transmission line. Point out both similarities and differences.

3. What is the relation between the wavelength in the cable and that in free space for a coaxial cable?

4. Can traveling waves with a continuous range of wavelengths be sent down (*a*) a coaxial cable and (*b*) a waveguide? Can standing waves with a continuous range of wavelengths be set up in a resonant cavity? Develop mechanical or acoustical analogies to support your answers.

5. If a certain wavelength is larger than the cutoff wavelength for a guide in its dominant mode, can energy be sent down it in any other mode?

6. Explain why the term $\epsilon_0\, d\Phi_E/dt$ is needed in Ampère's equation to understand the propagation of electromagnetic waves.

7. In the equation $c = 1/\sqrt{\mu_0\epsilon_0}$ (Eq. 39–15), how can c always have the same value if μ_0 is arbitrarily assigned and ϵ_0 is measured?

8. Is it conceivable that electromagnetic theory might some day be able to predict the value of c (3×10^8 meters/sec), not in terms of μ_0 and ϵ_0, but directly and numerically without recourse to any measurements?

9. What is the direction of the displacement current in Fig. 39–4? Give an argument to support your answer.

10. In a coaxial cable is the energy transported in the conductors, through the agency of the currents, or in the space between them, through the agency of the fields?

PROBLEMS

1. Using Gauss's law, sketch the instantaneous charges that appear on the conductors of the coaxial cable of Fig. 39–3 and show that this pattern of charges is appropriately related to the conduction currents shown in Fig. 39–3*c*.

2. For a rectangular guide of width 3.0 cm, plot the phase speed, the group speed, and the guide wavelength as a function of the free-space wavelength. Assume the dominant mode.

3. For a rectangular guide of width 3.0 cm, what must the free-space wavelength of radiation be if it is to require 1.0 μsec (= 10^{-6} sec) for energy to traverse a 100-meter length of guide? What is the phase speed under these circumstances?

4. Under what conditions will the guide wavelength in the guide of Fig. 39–7 be double the free-space wavelength?

5. How does the displacement current vary with space and time in a traveling plane electromagnetic wave?

6. Prove that for any point in an electromagnetic wave such as that of Fig. 39–11 the density of energy stored in the electric field equals that stored in the magnetic field.

7. A resonant cavity is constructed by closing each end of the coaxial cable of Fig. 39–3 with a metal cap. The cavity contains three half-waves. Describe the patterns of **E** and **B**

that occur, assuming the same mode of oscillation as that shown in Fig. 39–3. (Hint: Remember that **E** can have no tangential component at a conducting surface and that **B** and **E** must be 90° out of phase.)

8. If a coaxial cable has resistance, energy must flow from the fields into the conducting surfaces to provide the Joule heating. How must the electric lines of force of Fig. 39–3*a* be modified in this case? (Hint: The Poynting vector near the surface must have a component pointing toward the surface.)

9. What guide wavelength does 10-cm radiation (free-space wavelength) exhibit in a rectangular guide whose width is 6.0 cm? Assume the dominant mode. What is the cutoff wavelength for this guide?

10. Sketch five more figures to complete the sequence of Fig. 39–9, showing radiation from an oscillating dipole. Include an indication of the lines of **B** in your drawings.

11. Figure 39–15 shows a long resistanceless transmission line, delivering power from a battery to a resistive load. A steady current i exists as shown. (*a*) Sketch qualitatively the electric and magnetic fields around the line, and (*b*) show that, according to the Poynting vector point of view, energy travels from the battery to the resistor through the space around the line and not through the line itself. (Hint: Each conductor in the line is an equipotential surface, since the line has been assumed to have no resistance.)

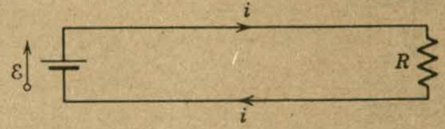


Fig. 39–15

12. Analyze the flow of energy in the waveguide of Fig. 39–7, using the Poynting vector.

13. A cube of edge a has its edges parallel to the x-, y-, and z-axes of a rectangular coordinate system.. A uniform electric field **E** is parallel to the y-axis and a uniform magnetic field **B** is parallel to the x-axis. Calculate (*a*) the rate at which, according to the Poynting vector point of view, energy may be said to pass through each face of the cube and (*b*) the net rate at which the energy stored in the cube may be said to change.

14. A #10 copper wire (diameter, 0.10 in.; resistance per 1000 ft, 1.00 ohm) carries a current of 25 amp. Calculate **E**, **B**, and **S** for a point on the surface of the wire.

15. Sunlight strikes the earth, outside its atmosphere, with an intensity of 2.0 cal/cm^2-min. Calculate E_m and B_m for sunlight, assuming it to be a wave like that of Fig. 39–10.

16. A plane radio wave has $E_m \cong 10^{-4}$ volt/meter. Calculate (*a*) B_m and (*b*) the intensity of the wave, as measured by $\overline{\mathbf{S}}$.

17. Figure 39–16 shows a cylindrical resistor of length l, radius a, and resistivity ρ, carrying a current i. (*a*) Show that the Poynting vector **S** at the surface of the resistor is everywhere directed normal to the surface, as shown. (*b*) Show that the rate P at which energy flows into the resistor through its cylindrical surface, calculated by integrating the Poynting vector over this surface, is equal to the rate at which Joule heat is produced; that is,

$$\int \mathbf{S}\cdot d\mathbf{A} = i^2R,$$

where $d\mathbf{A}$ is an element of area of the cylindrical surface. This shows that, according to the Poynting vector point of view, the energy that appears in a resistor as Joule heat does not enter it through the connecting wires but through the space around the wires and the resistor. (Hint: **E** is parallel to the axis of the cylinder, in the direction of the current; **B** forms concentric circles around the cylinder, in a direction given by the right-hand rule.)

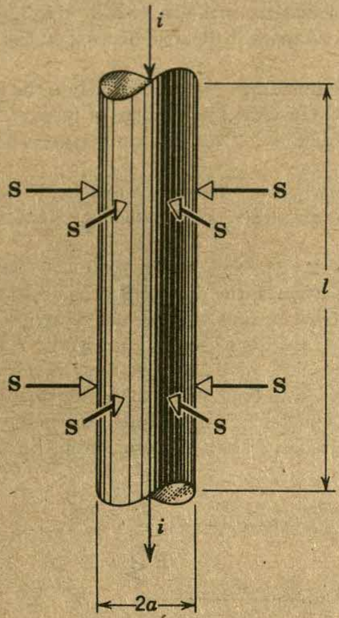


Fig. 39–16

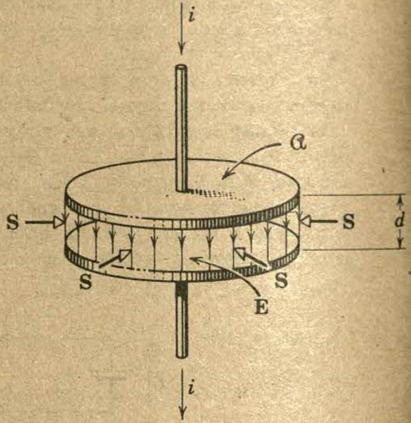


Fig. 39–17

18. Figure 39–17 shows a parallel-plate capacitor being charged. (*a*) Show that the Poynting vector **S** points everywhere radially into the cylindrical volume. (*b*) Show that the rate P at which energy flows into this volume, calculated by integrating the Poynting vector over the cylindrical boundary of this volume, is equal to the rate at which the stored electrostatic energy increases; that is, that

$$\int \mathbf{S} \cdot d\mathbf{A} = \mathcal{A} d \frac{d}{dt} \left(\tfrac{1}{2}\epsilon_0 E^2\right),$$

where $\mathcal{A}d$ is the volume of the capacitor and $\frac{1}{2}\epsilon_0 E^2$ is the energy density for all points within that volume. This analysis shows that, according to the Poynting vector point of view, the energy stored in a capacitor does not enter it through the wires but through the space around the wires and the plates. (Hint: To find **S** we must first find **B**, which is the magnetic field set up by the displacement current during the charging process; see Fig. 38–9. Ignore fringing of the lines of **E**.)

Nature and Propagation of Light

CHAPTER 40

40–1 Light and the Electromagnetic Spectrum

Light was shown by Maxwell to be a component of the *electromagnetic spectrum* of Fig. 40–1. All these waves are electromagnetic in nature and have the same speed c in free space. They differ in wavelength (and thus in frequency) only, which means that the sources that give rise to them and the instruments used to make measurements with them are rather different.* The electromagnetic spectrum has no definite upper or lower limit. The labeled regions in Fig. 40–1 represent frequency intervals within which a common body of experimental technique, such as common sources and common detectors, exists. All such regions overlap. For example, we can produce radiation of wavelength 10^{-3} meter either by microwave techniques (microwave oscillators) or by infrared techniques (incandescent sources).

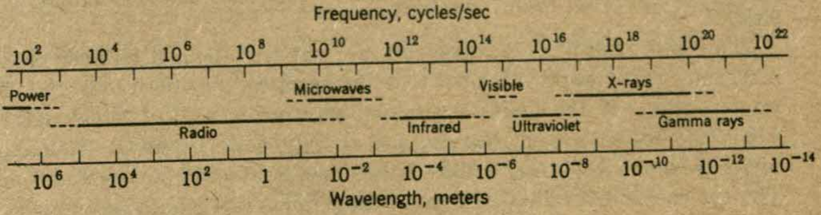


Fig. 40–1 The electromagnetic spectrum. Note that the wavelength and frequency scales are logarithmic.

---

* For a report of electromagnetic waves with wavelengths as long as 1.9×10^7 miles the student should consult an article by James Heirtzler in the *Scientific American* for March 1962.

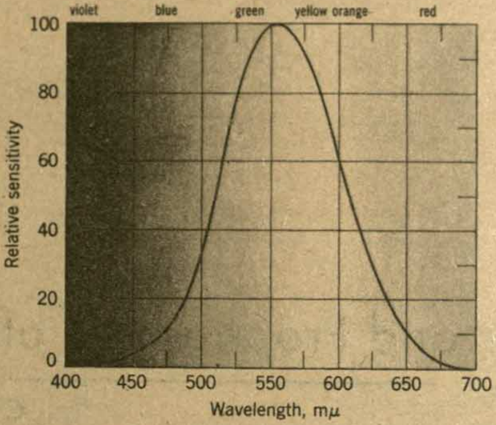


Fig. 40–2 The relative eye sensitivity of an assumed *standard observer* at different wavelengths for normal levels of illumination. The shaded areas represent the (continuously graded) color sensations for normal vision.

"Light" is defined here as radiation that can affect the eye. Figure 40–2, which shows the relative eye sensitivity of an assumed *standard observer* to radiations of various wavelengths, shows that the center of the visible region is about 5.55×10^{-7} meter. Light of this wavelength produces the sensation of yellow-green.*

In optics we often use the micron (abbr. μ) the millimicron (abbr. $m\mu$), and the Angstrom (abbr. A) as units of wavelength. They are defined from

$$1\ \mu = 10^{-6}\ \text{meter}$$

$$1\ m\mu = 10^{-9}\ \text{meter}$$

$$1\ \text{A} = 10^{-10}\ \text{meter.}$$

Thus the center of the visible region can be expressed as $0.555\ \mu$, $555\ m\mu$, or 5550 A.

The limits of the visible spectrum are not well defined because the eye sensitivity curve approaches the axis asymptotically at both long and short wavelengths. If the limits are taken, arbitrarily, as the wavelengths at which the eye sensitivity has dropped to 1% of its maximum value, these limits are about 4300 A and 6900 A, less than a factor of two in wavelength. The eye can detect radiation beyond these limits if it is intense enough. In many experiments in physics one can use photographic plates or light-sensitive electronic detectors in place of the human eye.

---

* See "Experiments in Color Vision" by Edwin H. Land, *Scientific American*, May 1959, and especially "Color and Perception: the Work of Edwin Land in the Light of Current Concepts" by M. H. Wilson and R. W. Brocklebank, *Contemporary Physics*, December 1961, for a fascinating discussion of the problems of perception and the distinction between color as a characteristic of light and color as a perceived property of objects.

40-2 Energy and Momentum

Energy is carried by electromagnetic waves from the sun to the earth or from an open fire to a hand placed nearby. The transport of energy by such a wave in free space was described in Section 39-6 by the Poynting vector **S**, or

$$\mathbf{S} = \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B}, \tag{40-1}$$

where **E** and **B** are the instantaneous values of the electric and magnetic field vectors.

Less familiar is the fact that electromagnetic waves may also transport linear momentum. In other words, it is possible to exert a pressure (a *radiation pressure* *) on an object by shining a light on it. Such forces must be small in relation to forces of our daily experience because we do not ordinarily notice them. The first measurement of radiation pressure was made in 1901–1903 by Nichols and Hull in this country and by Lebedev in Russia, about thirty years after the existence of such effects had been predicted theoretically by Maxwell.

Let a parallel beam of light fall on an object for a time t, the incident light being *entirely absorbed* by the object. If energy U is absorbed during this time, the momentum p delivered to the object is given, according to Maxwell's prediction, by

$$p = \frac{U}{c} \quad \text{(total absorption)}, \tag{40-2a}$$

where c is the speed of light. The direction of **p** is the direction of the incident beam. If the light energy U is *entirely reflected*, the momentum delivered will be twice that given above, or

$$p = \frac{2U}{c} \quad \text{(total reflection)}. \tag{40-2b}$$

In the same way, twice as much momentum is delivered to an object when a perfectly elastic tennis ball is bounced from it as when it is struck by a perfectly inelastic ball of the same mass and speed. If the light energy U is partly reflected and partly absorbed, the delivered momentum will lie between U/c and $2U/c$.

▶ **Example 1.** A parallel beam of light with an energy flux S of 10 watts/cm² falls for 1 hr on a perfectly reflecting plane mirror of 1.0-cm² area. (*a*) What momentum is delivered to the mirror in this time and (*b*) what force acts on the mirror?

(*a*) The energy that is reflected from the mirror is

$$U = (10 \text{ watts/cm}^2)(1.0 \text{ cm}^2)(3600 \text{ sec}) = 3.6 \times 10^4 \text{ joules}.$$

The momentum delivered after 1 hr's illumination is

$$p = \frac{2U}{c} = \frac{(2)(3.6 \times 10^4 \text{ joules})}{3 \times 10^8 \text{ meters/sec}} = 2.4 \times 10^{-4} \text{ kg-m/sec}.$$

---

* See "Radiation Pressure," G. E. Henry, *Scientific American*, p. 99, June 1957.

(*b*) From Newton's second law, the average force on the mirror is equal to th average rate at which momentum is delivered to the mirror, or

$$F = \frac{p}{t} = \frac{2.4 \times 10^{-4}\ \text{kg-m/sec}}{3600\ \text{sec}} = 6.7 \times 10^{-8}\ \text{nt}.$$

This is a small force. ◀

Nichols and Hull, in 1903, measured radiation pressures and verified Eq. 40–2, using a torsion balance technique. They allowed light to fall on mirror M in Fig. 40–3; the radiation pressure caused the balance arm to turn through a measured angle θ, twisting the torsion fiber F. Assuming a suitable calibration for their torsion fiber, the experimenters could arrive at a numerical value for this pressure. Nichols and Hull measured the intensity of their light beam by allowing it to fall on a blackened metal disk of known absorptivity and by measuring the temperature rise of this disk. In a particular run these experimenters measured a radiation pressure of 7.01×10^{-6} nt/meter^2; for their light beam, the value predicted, using Eq. 40–2, was 7.05×10^{-6} nt/meter^2, in excellent agreement. Assuming a mirror area of 1 cm^2, this represents a force on the mirror of only 7×10^{-10} nt, about 100 times smaller than the force calculated in Example 1.

The success of the experiment of Nichols and Hull was the result in large part of the care they took to eliminate spurious deflecting effects caused by changes in the speed distribution of the molecules in the gas surrounding the mirror. These changes were brought about by the small rise in the temperature of the mirror as it absorbed light energy from the incident beam. This "radiometer effect" is responsible for the spinning action of the familiar toy radiometers when placed in a beam of sunlight. In a perfect vacuum such effects would not occur, but in the best vacuums available in 1903 radiometer effects were present and had to be taken specifically into account in the design of the experiment.

To demonstrate the transport of momentum from Maxwell's equations in a particular case, let a plane electromagnetic wave traveling in the z direction fall on a

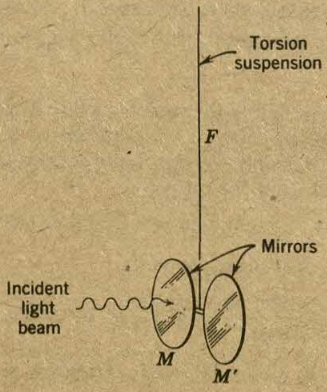


Fig. 40–3

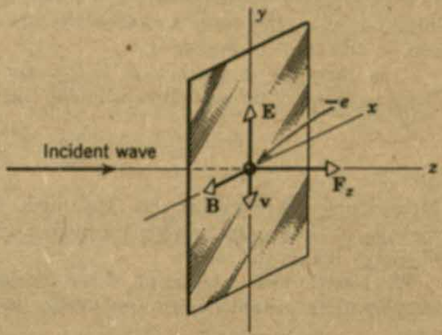


Fig. 40-4 An incident plane light wave falls on an electron in a thin resistive sheet. Instantaneous values of **E**, **B**, the electron velocity **v**, and the radiation force **F**$_z$ are shown.

large thin sheet of a material of high resistivity as in Fig. 40-4. A small part of the incident energy will be absorbed within the sheet, but most of it will be transmitted if the sheet is thin enough.*

The incident wave **E** and **B** vary with time at the sheet as

$$\mathbf{E} = \mathbf{E}_m \sin \omega t \tag{40-3}$$

and

$$\mathbf{B} = \mathbf{B}_m \sin \omega t \tag{40-4}$$

where **E** is parallel to the $\pm y$ axis and **B** is parallel to the $\pm x$ axis.

In Section 31-4 we saw that the effect of a (constant) electric force ($= -eE$) on a conduction electron in a metal was to make it move with a (constant) drift speed v_d. The electron behaves as if it is immersed in a viscous fluid, the electric force acting on it being counterbalanced by a "viscous" force, which may be taken as proportional to the electron speed. Thus for a constant field E, after equilibrium is established,

$$eE = bv_d, \tag{40-5}$$

where b is a resistive damping coefficient. The electron equilibrium speed, dropping the subscript d, is thus

$$v = \frac{eE}{b}. \tag{40-6}$$

If the applied electric field varies with time and if the variation is slow enough, the electron speed can continually readjust itself to the changing value of E so that its speed continues to be given essentially by its equilibrium value (Eq. 40-6) at all times. These readjustments are more rapidly made the more viscous the medium, just as a stone falling in air reaches a constant equilibrium rate of descent only relatively slowly but one falling in a viscous oil does so quite rapidly. We assume that the sheet in Fig. 40-4 is so viscous, that is, that its resistivity is so high, that Eq. 40-6 remains valid even for the rapid oscillations of E in the incident light beam.

As the electron vibrates parallel to the y axis, it experiences a *second* force due to the *magnetic* component of the wave. This force $\mathbf{F}_z$ ($= -e\mathbf{v} \times \mathbf{B}$) points in the z direction, being at right angles to the plane formed by **v** and **B**, that is, the xy plane. The instantaneous magnitude of $\mathbf{F}_z$ is given by

$$F_z = evB = \frac{e^2EB}{b}. \tag{40-7}$$

---

* Some of the incident energy will also be reflected, but the reflected wave is of such low intensity that it can be ignored in the derivation that follows; see *Optics* by B. Rossi, Addison-Wesley Publishing Company, p. 411, 1957, from which this derivation is adapted.

$\mathbf{F}_z$ always points in the positive z direction because $\mathbf{v}$ and $\mathbf{B}$ reverse their directions simultaneously; this force is, in fact, the mechanism by which the radiation pressure acts on the sheet of Fig. 40–4.

From Newton's second law, F_z is the rate dp_e/dt at which the incident wave delivers momentum to each electron in the sheet, or

$$\frac{dp_e}{dt} = \frac{e^2EB}{b}. \tag{40–8}$$

Momentum is delivered at this rate to every electron in the sheet and thus to the sheet itself. It remains to relate the momentum transfer to the sheet to the absorption of energy within the sheet.

The electric field component of the incident wave does work on each oscillating electron at an instantaneous rate (see Eq. 40–6) given by

$$\frac{dU_e}{dt} = F_E v = (eE)\left(\frac{eE}{b}\right) = \frac{e^2E^2}{b}.$$

Note that the magnetic force $\mathbf{F}_z$, always being at right angles to the velocity $\mathbf{v}$, does no work on the oscillating electron. Equation 39–11*b* shows that for a plane wave in free space B and E are related by

$$E = Bc.$$

Substituting above for *one* of the E's leads to

$$\frac{dU_e}{dt} = \frac{e^2EBc}{b}. \tag{40–9}$$

This equation represents the rate, per electron, at which energy is absorbed from the incident wave.

Comparing Eqs. 40–8 and 40–9 shows that

$$\frac{dp_e}{dt} = \frac{1}{c}\frac{dU_e}{dt}.$$

Integrating yields

$$\int_0^t \frac{dp_e}{dt}\,dt = \frac{1}{c}\int_0^t \frac{dU_e}{dt}\,dt,$$

or

$$p_e = \frac{U_e}{c}, \tag{40–10}$$

where p_e is the momentum delivered to a single electron in any given time t and U_e is the energy absorbed by that electron in the same time interval. Multiplying each side by the number of free electrons in the sheet leads to Eq. 40–2*a*.

Although we derived this relation (Eq. 40–10) for a particular kind of absorber, no characteristics of the absorber—for example, the resistive damping coefficient *b*—remain in the final expression. This is as it should be because Eq. 40–10 is a general property of radiation absorbed by *any* material.

40–3 The Speed of Light *

Light travels so fast that there is nothing in our daily experience to suggest that its speed is not infinite. It calls for considerable insight even to ask "How fast does light travel?" Galileo asked himself this question and tried to answer it experimentally. His chief work, *Two New Sciences*, published in the Netherlands in 1638, is written in the form of a conversation among three fictitious persons called Salviati, Sagredo, and Simplicio. Here is part of what they say about the speed of light.

---

* See "The Speed of Light," J. H. Rush, *Scientific American*, p. 67, August 1955.

> *Simplicio:* Everyday experience shows that the propagation of light is instantaneous; for when we see a piece of artillery fired, at a great distance, the flash reaches our eyes without lapse of time; but the sound reaches the ear only after a noticeable interval.
>
> *Sagredo:* Well, Simplicio, the only thing I am able to infer from this familiar bit of experience is that sound, in reaching our ear, travels more slowly than light; it does not inform me whether the coming of the light is instantaneous or whether, although extremely rapid, it still occupies time. . . .

Sagredo, who evidently is Galileo himself, then describes a possible method for measuring the speed of light. He and an assistant stand facing each other some distance apart, at night. Each carries a lantern which can be covered or uncovered at will. Galileo started the experiment by uncovering his lantern. When the light reached the assistant he uncovered his own lantern, whose light was then seen by Galileo. Galileo tried to measure the time between the instant at which he uncovered his own lantern and the instant at which the light from his assistant's lantern reached him. For a 1-mile separation we now know that the round trip travel time would be only 11×10^{-6} sec. This is much less than human reaction times, so the method fails.

To measure a large velocity directly, we must either measure a small time interval or use a long base line. This situation suggests that astronomy, which deals with great distances, might be able to provide an experimental value for the speed of light; this proved to be true. Although it would be desirable to time the light from the sun as it travels to the earth, there is no way of knowing when the light that reaches us at any instant left the sun; we must use subtler astronomical methods.

Note, however, that microwave pulses are quite regularly reflected from the moon; this gives a 7.68×10^8-meter base line (there and back) for timing purposes. The speed of light (and of microwaves) is so well known now from other experiments that these measurements are used to measure the lunar distance accurately. Microwave signals have also been reflected from Venus.

In 1675 Ole Roemer, a Danish astronomer working in Paris, made some observations of the moons of Jupiter (see Problem 9) from which a speed of light of 2×10^8 meters/sec may be deduced. About fifty years later James Bradley, an English astronomer, made some astronomical observations of an entirely different kind from which a value of 3.0×10^8 meters/sec may be deduced.

In 1849 Hippolyte Louis Fizeau (1819–1896), a French physicist, first measured the speed of light by a nonastronomical method, obtaining a value of 3.13×10^8 meters/sec. Figure 40–5 shows Fizeau's apparatus. Let us first ignore the toothed wheel. Light from source S is made to converge by lens L_1, is reflected from mirror M_1, and forms in space at F an image of the source. Mirror M_1 is a so-called "half-silvered mirror"; its reflecting coating is so thin that only half the light that falls on it is reflected, the other half being transmitted.

Light from the image at F enters lens L_2 and emerges as a parallel beam; after passing through lens L_3 it is reflected back along its original direction by mirror M_2. In Fizeau's experiment the distance l between M_2 and F was

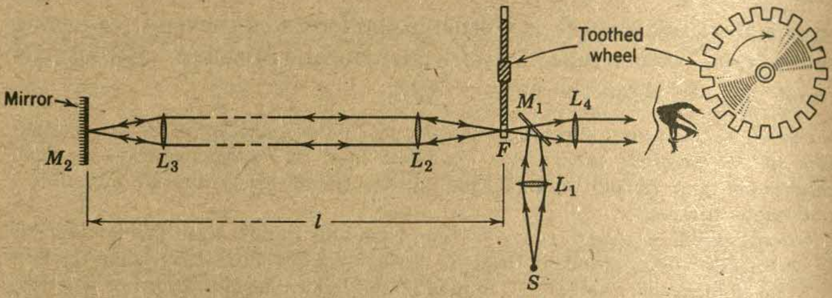


Fig. 40–5 Fizeau's apparatus for measuring the speed of light.

8630 meters or 5.36 miles. When the light strikes mirror M_1 again, some will be transmitted, entering the eye of the observer through lens L_4.

The observer will see an image of the source formed by light that has traveled a distance $2l$ between the wheel and mirror M_2 and back again. To time the light beam a marker of some sort must be put on it. This is done by "chopping" it with a rapidly rotating toothed wheel. Suppose that during the round-trip travel time of $2l/c$ the wheel has turned just enough so that, when the light from a given "burst" returns to the wheel, point F is covered by a tooth. The light will hit the face of the tooth that is toward M_2 and will not reach the observer's eye.

If the speed of the wheel is exactly right, the observer will not see any of the bursts because each will be screened by a tooth. The observer measures c by increasing the angular speed ω of the wheel from zero until the image of source S disappears. Let θ be the angular distance from the center of a gap to the center of a tooth. The time needed for the wheel to rotate a distance θ is the round-trip travel time $2l/c$. In equation form,

$$\frac{\theta}{\omega} = \frac{2l}{c} \quad \text{or} \quad c = \frac{2\omega l}{\theta}. \tag{40–11}$$

This "chopped beam" technique, suitably modified, is used today to measure the speeds of neutrons and other particles.

▶ **Example 2.** The wheel used by Fizeau had 720 teeth. What is the smallest angular speed at which the image of the source will vanish?

The angle θ is 1/1440 rev; solving Eq. 40–11 for ω gives

$$\omega = \frac{c\theta}{2l} = \frac{(3.00 \times 10^8 \text{ meters/sec})(1/1440 \text{ rev})}{(2)(8630 \text{ meters})} = 12.1 \text{ rev/sec.} \quad ◀$$

The French physicist Foucault (1819–1868) greatly improved Fizeau's method by substituting a rotating mirror for the toothed wheel. The American physicist Albert A. Michelson (1852–1931) conducted an extensive series of measurements of c, extending over a fifty-year period, using this technique.

Table 40–1

THE SPEED OF ELECTROMAGNETIC RADIATION IN FREE SPACE
(Some selected measurements)

Date	Experimenter	Country	Method	Speed, km/sec	Uncertainty, km/sec
1600(?)	Galileo	Italy	Lanterns and shutters	"If not instantaneous, it is extraordinarily rapid"	
1675	Roemer	France	Astronomical	200,000	
1729	Bradley	England	Astronomical	304,000	
1849	Fizeau	France	Toothed wheel	313,300	
1862	Foucault	France	Rotating mirror	298,000	500
1876	Cornu	France	Toothed wheel	299,990	200
1880	Michelson	U.S.A.	Rotating mirror	299,910	50
1883	Newcomb	England	Rotating mirror	299,860	30
1883	Michelson	U.S.A.	Rotating mirror	299,853	60
1906	Rosa and Dorsey	U.S.A.	Electromagnetic theory	299,781	10
1923	Mercier	France	Standing waves on wires	299,782	15
1926	Michelson	U.S.A.	Rotating mirror	299,796	4
1928	Karolus and Mittelstaedt	Germany	Kerr cell	299,778	10
1932	Michelson, Pease, and Pearson	U.S.A.	Rotating mirror	299,774	11
1940	Huettel	Germany	Kerr cell	299,768	10
1941	Anderson	U.S.A.	Kerr cell	299,776	14
1950	Bergstrand	Sweden	Geodimeter	299,792.7	0.25
1950	Essen	England	Microwave cavity	299,792.5	3
1950	Houston	Scotland	Vibrating crystal	299,775	9
1950	Bol and Hansen	U.S.A.	Microwave cavity	299,789.3	0.4
1951	Aslakson	U.S.A.	Shoran radar	299,794.2	1.9
1952	Rank, Ruth, and Ven der Sluis	U.S.A.	Molecular spectra	299,776	7
1952	Froome	England	Microwave interferometer	299,792.6	0.7
1954	Florman	U.S.A.	Radio interferometer	299,795.1	3.1
1954	Rank, Shearer, and Wiggins	U.S.A.	Molecular spectra	299,789.8	3.0
1956	Edge	Sweden	Geodimeter	299,792.9	0.2

We must view the speed of light within the larger framework of the speed of electromagnetic radiation in general. It is a significant experimental confirmation of Maxwell's theory of electromagnetism that the speed in free space of waves in all parts of the electromagnetic spectrum has the same value c. Table 40–1 shows some selected measurements that have been made of the speed of electromagnetic radiation since Galileo's day. It stands as a monument to man's persistence and ingenuity. Note in the last column how the uncertainty in the measurement has improved through the years. Note also the international character of the effort and the variety of methods.

The task of arriving at a single "best" value of c from the many listed in the table is difficult, for it involves a careful study of each reported measurement and a selection from among them, based on the reported uncertainty and the selector's judgment of the probable presence or absence of hidden error. In the final averaging measurements with small uncertainties will be given more weight than those with large ones. By careful analysis of such measurements the "best" value, as of 1964, is

$$c = 2.997925 \times 10^8 \text{ meters/sec.}$$

The uncertainty of measurement is less than 0.000003×10^8 meters/sec or 0.0001%.

Since about 1940, nearly all precise measurements of c have been made in the microwave or the short radio wave region of the electromagnetic spectrum. We describe here the "microwave cavity method" used by Essen in England and by Bol and Hansen in the U.S.A. It employs standing electromagnetic waves confined to a cavity rather than traveling waves in free space.

It is possible to convert a section of waveguide such as that of Fig. 39–7 into a resonant cavity by closing it with two metal caps; see Fig. 40–6. The pattern of oscillations in the cavity is closely related to that in the guide and exhibits the same "guide wavelength" λ_g. The guide wavelength is related to the cavity length l by

$$\lambda_g = \frac{2l}{n} \qquad n = 1, 2, 3, \cdots, \tag{40–12}$$

which is the same relationship used for acoustic waves in closed pipes; n (= 3 for Fig. 40–6) gives the number of half-waves contained in the cavity.

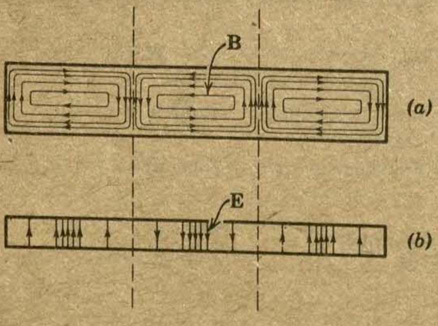


Fig. 40–6 A resonant cavity constructed from a section of waveguide; compare Fig. 39–7. For simplicity the lines of **E** are not shown in (*a*) and those of **B** are not shown in (*b*).

The procedure is to measure λ_g for such a cavity, which has been tuned to resonance, and then, using Eq. 39-4,

$$\lambda_g = \frac{\lambda}{\sqrt{1 - \left(\frac{\lambda}{2a}\right)^2}}, \tag{40-13}$$

calculate the free-space wavelength λ. From the measured resonant frequency, the speed c can be found from $c = \lambda\nu$.

▶ **Example 3.** Essen of the National Physical Laboratory in England made a resonant cavity measurement of the speed of electromagnetic waves. His cavity was made of a circular waveguide rather than a rectangular one; it can be shown that, for the oscillation pattern used by him, the geometrical factor $2a$ in Eq. 40-13 must be replaced by $1.64062R$, where R is the guide radius. The cavity radius was 3.25876 cm; the cavity length was 15.64574 cm and it proved to resonate at 9.498300×10^9 cycles/sec. At resonance it was determined that there were eight half-waves in the cavity. What value of c results?

From Eq. 40-12, computing only an approximate result,

$$\lambda_g = \frac{2l}{n} = \frac{(2)(15.6 \text{ cm})}{8} = 3.90 \text{ cm}.$$

Substituting into Eq. 40-13, suitably modified for a circular waveguide, yields

$$3.90 \text{ cm} = \frac{\lambda}{\sqrt{1 - \left(\frac{\lambda}{(1.64)(3.26 \text{ cm})}\right)^2}}.$$

Solving this equation for λ yields $\lambda = 3.15$ cm. Finally, we have

$$c = \lambda\nu = (3.15 \text{ cm})(9.50 \times 10^9 \text{ cycles/sec}) = 2.99 \times 10^8 \text{ meters/sec}.$$

For practical reasons Essen analyzed his data in a more roundabout way than that given. His final result, based on many measurements under different conditions and carried to much greater accuracy than that illustrated in the above example, was 2.997925×10^8 meters/sec with an uncertainty of 0.000030×10^8 meters/sec. ◀

40-4 Moving Sources and Observers

When we say that the speed of sound in dry air at 0°C is 331.7 meters/sec, we imply a reference frame fixed with respect to the air mass. When we say that the speed of light in free space is 2.997925×10^8 meters/sec, what reference frame is implied? It cannot be the medium through which the light wave travels because, in contrast to sound, no medium is required.

The concept of a wave requiring no medium was abhorrent to the physicists of the nineteenth century, influenced as they then were by a false analogy between light waves and sound waves or other purely mechanical disturbances. These physicists postulated the existence of an *ether*, which was a tenuous substance that filled all space and served as a medium of transmission for light. The ether was required to have a vanishingly small density to account for the fact that it could not be observed by any known means in an evacuated space.

The ether concept, although it proved useful for many years, did not survive the test of experiment. In particular, careful attempts to measure

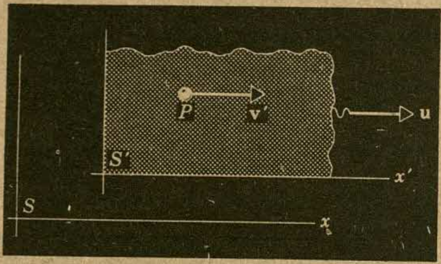


Fig. 40–7 Observers S and S', who are in relative motion, each observe a light pulse P. The pulse is emitted from a source, not shown, that is at rest in the S' frame of reference.

the speed of the earth through the ether always gave the result of zero.* Physicists were not willing to believe that the earth was permanently at rest in the ether and that all other bodies in the universe were in motion through it. Other hypotheses about the nature of the propagation of light also proved unsatisfactory for one reason or another.

Einstein in 1905 resolved the difficulty of understanding the propagation of light by making a bold postulate: If a number of observers are moving (at uniform velocity) with respect to each other and to a source of light and if each observer measures the speed of the light emerging from the source, *they will all obtain the same value.* This is the fundamental assumption of Einstein's theory of relativity. It does away with the need for an ether by asserting that the speed of light is the same in *all* reference frames; none is singled out as fundamental. The theory of relativity, derived from this postulate, has been subject to many experimental tests, from which agreement with the predictions of theory has always emerged. These agreements, extending over half a century, lend strong support to Einstein's basic postulate about light propagation.

Figure 40–7 focuses specifically on the fundamental problem of light propagation. A source of light, at rest in reference frame S', emits a light pulse P whose speed v' is measured by an observer at rest in this same frame. From the point of view of an observer in reference frame S, frame S' and its associated observer are moving in the positive x direction at speed u. Question: What speed v would observer S measure for the light pulse P? Einstein's hypothesis asserts that *each* observer would measure the same speed c, or that

$$v = v' = c.$$

This hypothesis contradicts the classical law of addition of velocities (see Section 4–6), which asserts that

$$v = v' + u. \tag{40-14}$$

This law, which is so familiar that it seems (incorrectly) to be intuitively true, is in fact based on observations of gross moving objects in the world about us. Even the fastest of these—an earth satellite, say—is moving at a speed that is quite small compared to that of light. The body of experimental evidence that underlies Eq. 40–14 thus represents a severely restricted area of experience, namely, experiences in which $v' \ll c$ and $u \ll c$. If we assume that Eq. 40–14 holds for all particles regardless of speed, we are making a gross extrapolation. Einstein's theory of relativity predicts that this extrapolation is indeed not valid and that Eq. 40–14 is a limiting case of a more general relationship that holds for light pulses and for material particles, whatever their speed, or

---

* See Section 43–7, which describes the crucial experiment of Michelson and Morley.

$$v = \frac{v' + u}{1 + v'u/c^2}. \tag{40–15}$$

Equation 40–15 is quite indistinguishable from Eq. 40–14 at low speeds that is, when $v' \ll c$ and $u \ll c$; see Example 4.

If we apply Eq. 40–15 to the case in which the moving object is a ligh pulse, and if we put $v' = c$, we obtain

$$v = \frac{c + u}{1 + cu/c^2} = c.$$

This is consistent, as it must be, with the fundamental assumption on whic the derivation of Eq. 40–15 is based; it shows that *both* observers measur the same speed c for light. Equation 40–14 predicts (incorrectly) that th speed measured by S will be $c + u$. Figure 40–8 shows that the (correct) Eq. 40–15 and the (approximate) Eq. 40–14 cannot be distinguished from each other at speeds that are small compared to the speed of light.

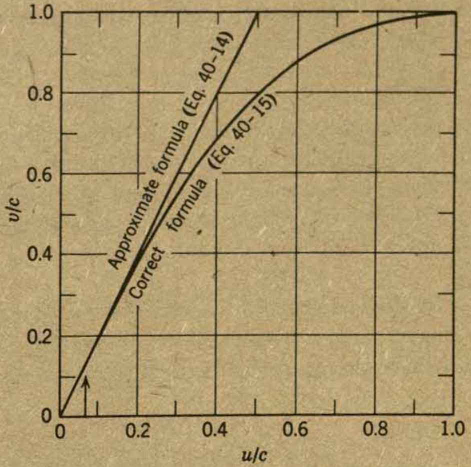


Fig. 40–8 The speed of a particle P, as seen by observer S in Fig. 40–7, for the special case of $v' = u$. All speeds are expressed as a ratio to c, the speed of light. The vertical arrow corresponds to 5×10^7 miles/hr.

▶ **Example 4.** Suppose that $v' = u = 25{,}000$ miles/hr. What per cent error is made in using Eq. 40–14 rather than Eq. 40–15 to calculate v? The speed of light is 6.7×10^8 miles/hr.

Equation 40–14 gives

$$v = v' + u = 25{,}000 \text{ miles/hr} + 25{,}000 \text{ miles/hr} = 50{,}000 \text{ miles/hr}.$$

Equation 40–15 gives

$$\begin{aligned} v &= \frac{v' + u}{1 + v'u/c^2} \\ &= \frac{25{,}000 \text{ miles/hr} + 25{,}000 \text{ miles/hr}}{1 + \dfrac{(25{,}000 \text{ miles/hr})^2}{(6.7 \times 10^8 \text{ miles/hr})^2}} \\ &= \frac{50{,}000 \text{ miles/hr}}{1.0000000014}. \end{aligned}$$

Even at 25,000 miles/hr the error in Eq. 40–14 is immeasurably small.

Example 5. Two electrons are ejected in opposite directions from radioactive atoms in a sample of radioactive material. Let each electron have a speed, as measured by a laboratory observer, of $0.6c$ (this corresponds to a kinetic energy of 130 kev). What is the speed of one electron as seen from the other?

Equation 40–14 gives

$$v = v' + u = 0.6c + 0.6c = 1.2c.$$

Equation 40–15 gives

$$v = \frac{v' + u}{1 + v'u/c^2} = \frac{0.6c + 0.6c}{1 + (0.6c)^2/c^2} = 0.88c.$$

This example shows that for speeds that are comparable to c, Eqs. 40–14 and 40–15 yield rather different results. A wealth of indirect experimental evidence points to the latter result as being correct. ◀

40–5 Doppler Effect

We have seen that the same speed is measured for light no matter what the relative speeds of the light source and the observer are. The measured frequency and wavelength will change, but always in such a way that their product, which is the velocity of light, remains constant. Such frequency shifts are called *Doppler shifts*, after Johann Doppler (1803–1853), who first predicted them.

In Section 20–7 we showed that if a source of *sound* is moving away from an observer at a speed u, the frequency heard by the observer (see Eq. 20–10, which has been rearranged with u substituted for v_s) is

$$\nu' = \nu \frac{1}{1 + u/v}. \quad \left\{ \begin{array}{l} \text{1. sound wave} \\ \text{2. observer fixed in medium} \\ \text{3. source receding from observer} \end{array} \right. \qquad (40\text{–}16)$$

In this equation ν is the frequency heard when the source is at rest and v is the speed of sound.

If the source is at rest in the transmitting medium but the observer is moving away from the source at speed u, the observed frequency (see Eq. 20-9b, in which u has been substituted for v_0) is

$$\nu' = \nu\left(1 - \frac{u}{v}\right). \qquad \begin{cases} \text{1. sound wave} \\ \text{2. source fixed in medium} \\ \text{3. observer receding from source} \end{cases} \qquad (40\text{-}17)$$

Even if the relative separation speeds u of the source and the observer are the same, the frequencies predicted by Eqs. 40-16 and 40-17 are different. This is not surprising, because a sound source moving through a medium in which the observer is at rest is physically different from an observer moving through that medium with the source at rest, as comparison of Figs. 20-10 and 20-11 shows.

We might be tempted to apply Eqs. 40-16 and 40-17 to light, substituting c, the speed of light, for v, the speed of sound. For light, as contrasted with sound, however, it has proved impossible to identify a medium of transmission relative to which the source and the observer are moving. This means that "source receding from observer" and "observer receding from source" are physically identical situations and must exhibit *exactly the same* Doppler frequency. As applied to light, either Eq. 40-16 or Eq. 40-17 or both must be incorrect. The Doppler frequency predicted by the theory of relativity is, in fact,

$$\nu' = \nu\frac{1 - u/c}{\sqrt{1 - (u/c)^2}}. \qquad \begin{cases} \text{1. light wave} \\ \text{2. source and observer} \\ \quad \text{separating} \end{cases} \qquad (40\text{-}18)$$

In all three of the foregoing equations we obtain the appropriate relations for the source and the observer *approaching* each other if we replace u by $-u$.

Equations 40-16, 40-17, and 40-18 are not so different as they seem if the ratio u/c is small enough. This was made clear in Example 3, Chapter 20, for the first two of these equations. Let us expand Eqs. 40-16, 40-17 and 40-18 by the binomial theorem, as in the example referred to. The equations then become, substituting c for v,

$$\nu' = \nu\left[1 - \frac{u}{c} + \left(\frac{u}{c}\right)^2 + \cdots\right], \qquad (40\text{-}16a)$$

$$\nu' = \nu\left(1 - \frac{u}{c}\right), \qquad (40\text{-}17a)$$

and

$$\nu' = \nu\left[1 - \frac{u}{c} + \frac{1}{2}\left(\frac{u}{c}\right)^2 + \cdots\right]. \qquad (40\text{-}18a)$$

The ratio u/c for all available monochromatic light sources, even those of atomic dimensions, is small. This means that successive terms in these equations become small rapidly and, depending on the accuracy required, only a limited number of terms need be retained.

Under nearly all circumstances the differences among these three equations are not important. Nevertheless, it is of extreme interest to carry out at least one experiment precisely enough to serve as a test of Eq. 40–18*a* and thus, in part, of the theory of relativity.

H. E. Ives and G. R. Stilwell carried out such a precision experiment in 1938. They sent a beam of hydrogen atoms, generated in a gas discharge, down a tube at speed *u*, as in Fig. 40–9*a*. They could observe light emitted by these atoms in a direction opposite to **u** (atom 1, for example) using a mirror, and also in a direction parallel to **u** (atom 2, for example). With a precision spectrograph, they could photograph a particular characteristic spectrum line in this light, obtaining, on a frequency scale, the lines marked ν_1' and ν_2' in Fig. 40–9*b*. It is also possible to photograph, on the same photographic plate, a line corresponding to light emitted from *resting* atoms; such a line appears as ν in Fig. 40–9*b*. A fundamental measured quantity in this experiment is $\Delta\nu/\nu$, defined from

$$\frac{\Delta\nu}{\nu} = \frac{\Delta\nu_2 - \Delta\nu_1}{\nu}, \tag{40–19}$$

(see Fig. 40–9*b*). It measures the extent to which the frequency of the light from resting atoms fails to lie halfway between the frequencies ν_1' and ν_2'. Table 40–2 shows that the measured results agree with the formula predicted by the theory of relativity (Eq. 40–18*a*) and not with the classical formula borrowed from the theory of sound propagation in a material medium (Eq. 40–16*a*).

Ives and Stilwell did not present their experimental results as evidence for the support of Einstein's theory of relativity but rather gave them an alternative theoretical explanation. Modern observers, looking not only at their excellent experiment but at the whole range of experimental evidence, now give the Ives-Stilwell experiment the interpretation we have described for it above.

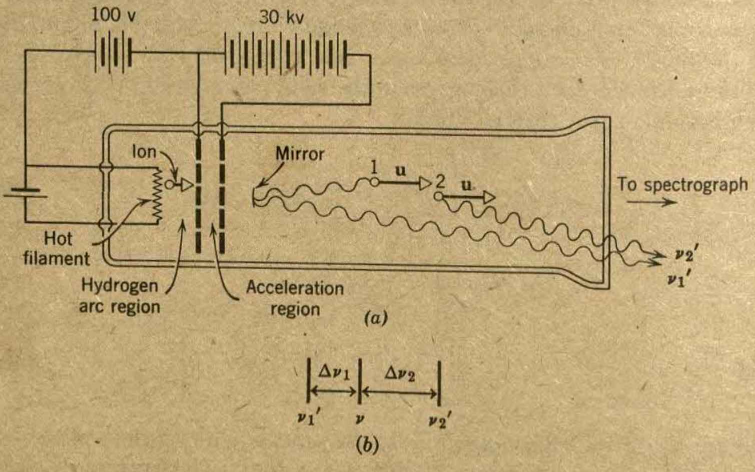


Fig. 40–9 The Ives-Stilwell experiment.

Table 40–2

THE IVES-STILWELL EXPERIMENT *

$\frac{\Delta\nu}{\nu}$, 10^{-5}	Speed of moving atoms ($= u$), 10^6 meters/sec			
	0.865	1.01	1.15	1.33
Theoretical value according to classical theory (Eq. 40–16*a*)	1.67	2.26	2.90	3.94
Theoretical value according to the theory of relativity (Eq. 40–18*a*)	0.835	1.13	1.45	1.97
Experimental value	0.762	1.1	1.42	1.9

* See Eq. 40–19; the table shows only part of the data taken by Ives and Stilwell.

The Doppler effect for light finds many applications in astronomy, where it is used to determine the speeds at which luminous heavenly bodies are moving toward us or receding from us. Such Doppler shifts measure only the radial or line-of-sight components of the relative velocity. All galaxies * for which such measurements have been made (Fig. 40–10) appear to be receding from us, the recession velocity being greater for the more distant galaxies; these observations are the basis of the expanding-universe concept.

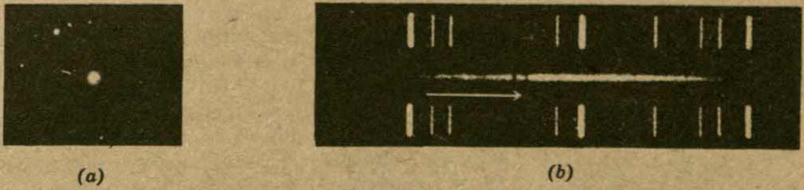

(*a*) (*b*)

Fig. 40–10 (*a*) The central spot is a nebula in the constellation Corona Borealis; it is 130,000,000 light years distant. (*b*) The central streak shows the distribution in wavelength of the light emitted from this nebula. The two vertical dark bands show the presence of calcium. The horizontal arrow shows that these calcium lines occur at longer wavelengths than those for terrestrial light sources containing calcium, the length of the arrow representing the wavelength shift. Measurement of this shift shows that the galaxy is receding from us at 13,400 miles/sec. The lines above and below the central streak represent light from a terrestrial source, used to establish a wavelength scale. (Courtesy Mount Wilson and Mount Palomar Observatories.)

---

* See "The Red-Shift," Allen R. Sandage, *Scientific American*, p. 171, September 1956.

▶ **Example 6.** Certain characteristic wavelengths in the light from a galaxy in the constellation Virgo are observed to be increased in wavelength, as compared with terrestrial sources, by about 0.4%. What is the radial speed of this galaxy with respect to the earth? Is it approaching or receding?

If λ is the wavelength for a terrestrial source, then

$$\lambda' = 1.004\lambda.$$

Since we must have $\lambda'\nu' = \lambda\nu = c$, we can write this as

$$\nu' = 0.996\nu.$$

This frequency shift is so small that, in calculating the source velocity, it makes no practical difference whether we use Eq. 40–16, 40–17, or 40–18. Using Eq. 40–17 we obtain

$$\nu' = 0.996\nu = \nu\left(1 - \frac{u}{c}\right).$$

Solving yields $u/c \cong 0.004$, or $u = (0.004)(3 \times 10^8 \text{ meters/sec}) = 1.2 \times 10^6$ meters/sec or 2.7×10^6 miles/hr. The galaxy is *receding;* had u turned out to be negative, the galaxy would have been moving toward us. ◀

QUESTIONS

1. How might an eye-sensitivity curve like that of Fig. 40–2 be measured?

2. Why are danger signals in red, when the eye is most sensitive to yellow-green?

3. Comment on this definition of the limits of the spectrum of visible light given by a physiologist: "The limits of the visible spectrum occur when the eye is no better adapted than any other organ of the body to serve as a detector."

4. How can an object absorb light energy without absorbing momentum?

5. A searchlight sends out a parallel beam of light. Does the searchlight experience any force associated with the emission of light?

6. Name two historic experiments, in addition to the radiation pressure measurements of Nichols and Hull, in which a torsion balance was used. Both are described in this book, one in Part 1 and one in Part 2.

7. Show that for complete absorption of a parallel beam of light the radiation pressure on the absorbing object is given by $p = S/c$, where S is the magnitude of the Poynting vector and c is the speed of light in free space.

8. How could Galileo test experimentally that reaction times were an overwhelming source of error in his attempt to measure the speed of light, described on p. 999?

9. It has been suggested that the velocity of light may change slightly in value as time goes on. Can you find any evidence for this in Table 40–1?

10. A friend asserts that Einstein's postulate (that the speed of light is not affected by the uniform motion of the source or the observer) must be discarded because it violates "common sense." How would you answer him?

11. In a vacuum, does the speed of light depend on (*a*) the wavelength, (*b*) the frequency, (*c*) the intensity, (*d*) the speed of the source, or (*e*) the speed of the observer?

12. Can a galaxy be so distant that its recession speed equals c? If so, how can we see the galaxy? That is, will its light ever reach us?

PROBLEMS

1. (*a*) At what wavelengths does the eye sensitivity have half its maximum value? (*b*) What are the frequency and the period of the light for which the eye is most sensitive?

2. It has been proposed that a spaceship might propel itself in the solar system by radiation pressure, using a large sail made of aluminum foil. How large must the sail be if the radiation force is to be equal in magnitude to the sun's gravitational attraction? Assume that the mass of the *ship* + *sail* is 100 slugs, that the sail is perfectly reflecting, and that the sail is oriented at right angles to the sun's rays. The sun's mass is 1.97×10^{30} kg.

3. Radiation from the sun striking the earth has an intensity of 1400 watts/meter^2. Assuming that the earth behaves like a flat disk at right angles to the sun's rays and that all the incident energy is absorbed, calculate the force on the earth due to radiation pressure. Compare it with the force due to the sun's gravitational attraction.

4. Prove, for a plane wave at normal incidence on a plane surface, that the radiation pressure on the surface is equal to the energy density in the beam outside the surface. This relation holds no matter what fraction of the incident energy is reflected.

5. Prove, for a stream of bullets striking a plane surface at right angles, that the "pressure" is *twice* the (kinetic) energy density in the stream above the surface; assume that the bullets are completely "absorbed" by the surface. Contrast this with the behavior of light (Problem 4).

6. A small spaceship whose mass, with occupant, is 100 slugs is drifting in outer space, where no gravitational field exists. If it shines a searchlight, which radiates 10^4 watts, into space, what speed would the ship attain in one day because of the reaction force associated with the momentum carried away by the light beam?

7. What is the radiation pressure 1.0 meter away from a 500-watt light bulb? Assume that the surface on which the pressure is exerted faces the bulb and is perfectly absorbing and that the bulb radiates uniformly in all directions.

8. The uncertainty of the distance to the moon, as measured by the reflection of radar waves from it, is about 0.5 mile. Assuming that this uncertainty is associated only with the measurement of the elapsed time, what uncertainty in this time is implied?

9. Roemer's method for measuring the speed of light consisted in observing the apparent times of revolution of one of the moons of Jupiter. The true period of revolution is 42.5 hr. (*a*) Taking into account the finite speed of light, how would you expect the apparent time of revolution to alter as the earth moves in its orbit from point x to point y in Fig. 40–11? (*b*) What observations would be needed to compute the speed of light? Neglect the motion of Jupiter in its orbit. Figure 40–11 is not drawn to scale.

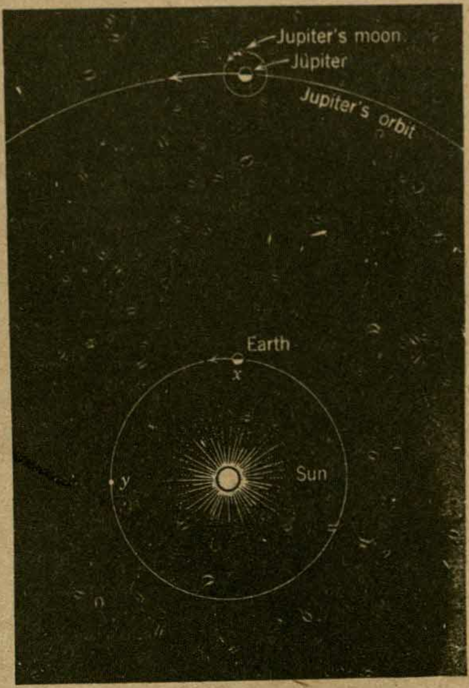


Fig. 40–11

10. Suppose that light is timed over a 1-mile base line and its speed is measured to the accuracy quoted on p. 1002. How large an error in the length of the base line could be tolerated, assuming other sources of error to be negligible?

11. For what value of u/c does Eq. 40–17 differ from Eq. 40–18 by 1%?

12. The "red shift" of radiation from a distant nebula consists of the light (H_γ), known to have a wavelength of 4340×10^{-8} cm when observed in the laboratory, appearing to have a wavelength of 6562×10^{-8} cm. What is the speed of the nebula in the line of sight relative to the earth? (*b*) Is it approaching or receding?

13. The difference in wavelength between an incident microwave beam and one reflected from an approaching or receding car is used to determine automobile speeds on the highway. (*a*) Show that if v is the speed of the car and ν the frequency of the incident beam, the change of frequency is approximately $2v\nu/c$, where c is the speed of the electromagnetic radiation. (*b*) For microwaves of frequency 2450 megacycles/sec, what is the change of frequency per mile/hr of speed?

14. Show that, for slow speeds, the Doppler shift can be written in the approximate form

$$\frac{\Delta\lambda}{\lambda} = \frac{u}{c},$$

where $\Delta\lambda$ is the change in wavelength.

15. The period of rotation of the sun at its equator is 24.7 days; its radius is 7.0×10^8 meters. What Doppler wavelength shifts are expected for characteristic wavelengths in the vicinity of 5500 A emitted from the edge of the sun's disk?

16. An earth satellite, transmitting on a frequency of 40×10^6 cycles/sec (exactly), passes directly over a radio receiving station at an altitude of 250 miles and at a speed of 18,000 miles/hr. Plot the change in frequency, attributable to the Doppler effect, as a function of time, counting $t = 0$ as the instant the satellite is over the station. (Hint: The speed u in the Doppler formula is not the actual velocity of the satellite but its component in the direction of the station. Use the nonrelativistic formula (Eq. 40–16*a*) and neglect the curvature of the earth and of the satellite orbit.)

17. A rocketship is receding from the earth at a speed of $0.2c$. A light in the rocketship appears blue to passengers on the ship. What color would it appear to be to an observer on the earth? See Fig. 40–2.

18. In the experiment of Ives and Stilwell the speed u of the hydrogen atoms in a particular run was 8.61×10^5 meters/sec. Calculate $\Delta\nu_1$, $\Delta\nu_2$, and $\Delta\nu/\nu$, on the assumptions that (*a*) Eq. 40–18*a* is correct and (*b*) that Eq. 40–16*a* is correct; compare your results with those given in Table 40–2 for this speed. Retain the first three terms only in Eqs. 40–18*a* and 40–16*a*.

Reflection and Refraction—Plane Waves and Plane Surfaces

CHAPTER 41

41-1 Reflection and Refraction

In Fig. 41-1*a* a light beam falling on a water surface is both reflected from the surface and bent (that is, *refracted*) as it enters the water. The incident beam is represented in Fig. 41-1*b* by a single line, the *incident ray*, parallel to the direction of propagation. The incident beam is assumed in Fig. 41-1*b* to be a *plane wave*, the wavefronts being normal to the incident ray. The reflected and refracted beams are also represented by rays. The angles of *incidence* (θ_1), of *reflection* (θ_1'), and of *refraction* (θ_2) are measured between the normal to the surface and the appropriate ray, as shown in the figure.

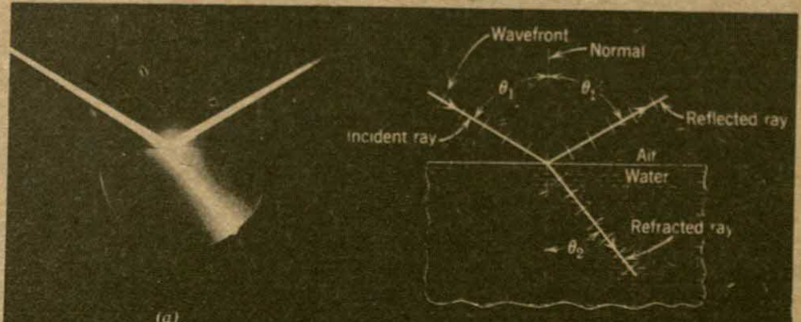


Fig. 41-1 (*a*) A photograph showing reflection and refraction at an air-water interface. (*b*) A representation using rays.

The laws governing reflection and refraction can easily be found from experiment:

1. The reflected and the refracted rays lie in the plane formed by the incident ray and the normal to the surface at the point of incidence, that is, the plane of Fig. 41–1*b*.

2. For reflection:

$$\theta_1' = \theta_1. \tag{41-1}$$

3. For refraction:

$$\frac{\sin \theta_1}{\sin \theta_2} = n_{21}, \tag{41-2}$$

where n_{21} is a constant called the *index of refraction* of medium 2 with respect to medium 1. Table 41–1 shows the indices of refraction for some common substances with respect to a vacuum for a wavelength of 5890 A (sodium light).

Table 41–1

SOME INDICES OF REFRACTION *
(For λ = 5890 A)

Medium	Index of Refraction
Water	1.33
Ethyl alcohol	1.36
Carbon bisulfide	1.63
Air (1 atm and 20°C)	1.0003
Methylene iodide	1.74
Fused quartz	1.46
Glass, crown	1.52
Glass, dense flint	1.66
Sodium chloride	1.53
Polyethylene	1.50–1.54
Fluorite	1.43

* Measured with respect to a vacuum. The index with respect to air (except, of course, the index of air itself—see item 4) will be negligibly different in most cases.

The index of refraction of one medium with respect to another generally varies with wavelength, as Fig. 41–2 shows. Because of this fact refraction, unlike reflection, can be used to analyze a beam of light into its component wavelengths. Figure 41–3, taken from Newton's *Opticks*, shows how Newton, using glass prism *ABC*, formed a spectrum of sunlight entering his window through a small hole at *F*.

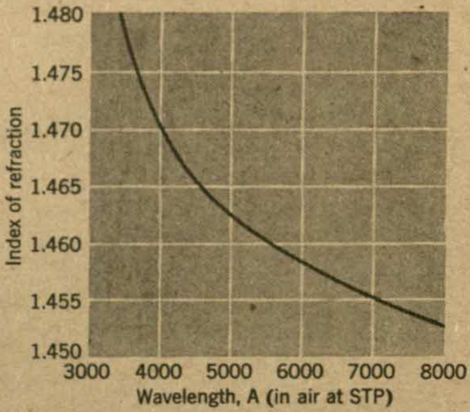


Fig. 41-2 The index of refraction of fused quartz at 18°C, with respect to a vacuum.

The law of reflection was known to Euclid. That of refraction was discovered experimentally by Willebrod Snell (1591–1626) and deduced from the early corpuscular theory of light by René Descartes (1596–1650). The law of refraction is known as Snell's law or (in France) as Descartes' law.

The laws of reflection and refraction can be derived from Maxwell's equations, which means that these laws should hold for all regions of the electromagnetic spectrum. Figure 41–4*a* shows an experimental setup for investigating the reflection of microwaves from a large metal sheet. Figure 41–4*b* shows the reading of the detector as a function of the angular position of the mirror. The existence of a reflected beam at the proper angle confirms the law of reflection for microwaves. There is ample experimental evidence that Eqs. 41–1 and 41–2 correctly describe the behavior of reflected and refracted beams in all parts of the electromagnetic spectrum.

It is common knowledge that a polished steel surface will form a well-defined reflected beam if an incident beam falls on it, but a sheet of paper will reflect light more or less in all directions (*diffuse reflection*). It is largely by diffuse reflection that we see nonluminous objects around us. The difference between diffuse and *specular* (that is, mirror-like) reflection is a matter of surface roughness; a reflected beam will

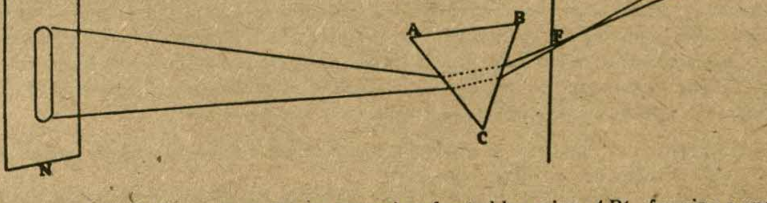


Fig. 41-3 Sunlight from a hole F in a screen is refracted by prism ABC, forming a spectrum on screen MN; from Newton's *Opticks* (1704).

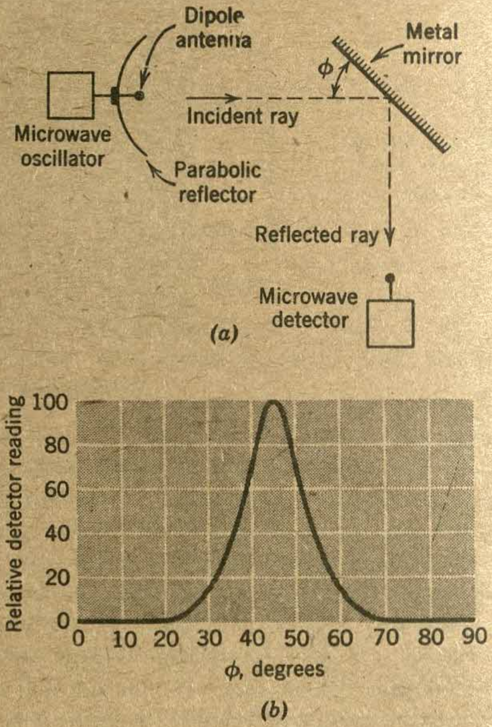


Fig. 41–4 (*a*) An apparatus to test the law of reflection for microwaves. (*b*) A reflected beam (for $\lambda \cong 10$ cm) appears for the expected orientation of the mirror.

be formed only if the average depth of the surface irregularities of the reflector is substantially less than the wavelength of the incident light. This criterion of surface roughness has different implications in different regions of the electromagnetic spectrum. The bottom of a cast-iron skillet for example is a good reflector for microwaves of wavelength 0.5 cm but it is not a good reflector for visible light (that is, one cannot shave by it).

A second requirement for the existence of a reflected beam is that the transverse dimensions of the reflector must be substantially larger than the wavelength of the incident beam. If a beam of visible light falls on a polished metal disk the size of a dime, a reflected beam will be formed. However, if the same disk is placed in a beam of short radio waves with, say, $\lambda = 1.0$ meter, radiation will be scattered in all directions from it, and no well-defined unidirectional beam will appear. We investigate this phenomenon of *diffraction* in Chapter 44. The requirements that surfaces be "smooth" and "large" also apply to the formation of refracted beams. If these two requirements are not met the description of reflection and refraction in terms of rays, whose behavior is governed by Eqs. 41–1 and 41–2, is not valid.

▶ **Example 1.** Figure 41–5 shows an incident ray i striking a plane mirror MM' at angle of incidence θ. Trace this ray.

The reflected ray makes an angle θ with the normal at b and falls as an incident ray on mirror $M'M''$. Its angle of incidence θ' on this mirror is $\pi/2 - \theta$. A second reflected ray r' makes an angle θ' with the normal erected at b'. Rays i and r' are antiparallel for any value of θ. To see this, note that

$$\phi = \pi - 2\theta' = \pi - 2\left(\frac{\pi}{2} - \theta\right) = 2\theta.$$

Two lines are parallel if their opposite interior angles for an intersecting line (ϕ and 2θ) are equal.

Repeat the problem if the angle between the mirrors is 120° rather than 90°.

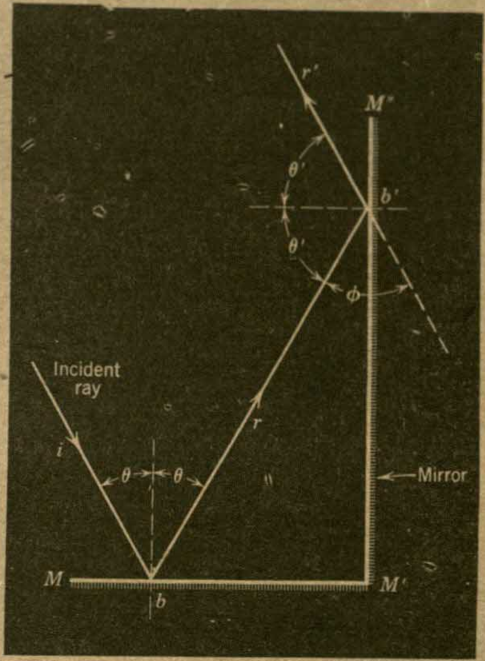


Fig. 41-5 Example 1.

Example 2. An incident ray in air falls on the plane surface of a block of quartz and makes an angle of 30° with the normal. This beam contains two wavelengths, 4000 and 5000 A. The indices of refraction for quartz with respect to air (n_{qa}) at these wavelengths are 1.4702 and 1.4624, respectively. What is the angle between the two refracted beams?

From Eq. 41-2 we have. for the 4000-A beam,

$$\sin\theta_1 = n_{qa}\sin\theta_2,$$

or

$$\sin 30^\circ = (1.4702)\sin\theta_2,$$

which leads to

$$\theta_2 = 19.88^\circ.$$

For the 5000-A beam we have

$$\sin 30^\circ = (1.4624)\sin\theta_2',$$

or

$$\theta_2' = 19.99^\circ.$$

The angle $\Delta\theta$ between the beams is 0.11°, the shorter wavelength component being bent through the larger angle, that is, having the smaller angle of refraction.

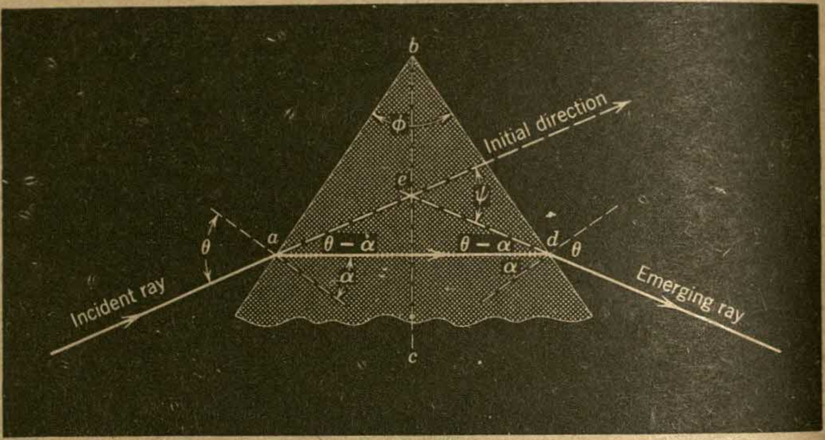


Fig. 41–6 Example 3.

Example 3. An incident ray falls on one face of a glass prism in air as in Fig. 41–6. The angle θ is so chosen that the emerging ray also makes an angle θ with the normal to the other face. Derive an expression for the index of refraction of the prism material with respect to air.

Note that $\angle\, abc = \alpha$, the two angles having their sides mutually perpendicular. Therefore

$$\alpha = \tfrac{1}{2}\phi. \qquad (41\text{–}3)$$

The *deviation angle* ψ is the sum of the two opposite interior angles in triangle *aed*, or

$$\psi = 2(\theta - \alpha).$$

Substituting $\frac{1}{2}\phi$ for α and solving for θ yields

$$\theta = \tfrac{1}{2}(\psi + \phi). \qquad (41\text{–}4)$$

At point a, θ is the angle of incidence and α the angle of refraction. The law of refraction (see Eq. 41–2) is

$$\sin\theta = n_{ga}\sin\alpha,$$

in which n_{ga} is the index of refraction of the glass with respect to air.

From Eqs. 41–3 and 41–4 this yields

$$\sin\frac{\psi + \phi}{2} = n_{ga}\sin\frac{\phi}{2}$$

or

$$n_{ga} = \frac{\sin\frac{1}{2}(\psi + \phi)}{\sin(\phi/2)},$$

which is the desired relation. This equation holds only for θ so chosen that the light ray passes symmetrically through the prism. For this condition the deviation angle ψ is a minimum; if θ is either increased or decreased, a larger deviation will be produced. ◀

41–2 Huygens' Principle

A theory of light would not be accepted if it were not able to predict the well-established laws of reflection and refraction. These laws can be derived from Maxwell's equations, but mathematical complexity prevents us from doing so. Fortunately, these and several other laws of optics can be derived on the basis of a simpler but less comprehensive theory of light, put forward by the Dutch physicist Christian Huygens in 1678. This theory simply assumes that light is a wave rather than, say, a stream of particles. It says nothing about the nature of the wave and, in particular—since Maxwell's theory of electromagnetism appeared only after the lapse of a century—gives no hint of the electromagnetic character of light. Huygens did not know whether light was a transverse wave or a longitudinal one; he did not know the wavelengths of visible light; he had little knowledge of the speed of light. Nevertheless, his theory was a useful guide to experiment for many years and remains useful today for pedagogic and certain other practical purposes. We must not expect it to yield the same wealth of detailed information that Maxwell's more complete electromagnetic theory does.

Huygens' theory is based on a geometrical construction, called *Huygens' principle*, that allows us to tell where a given wavefront will be at any time in the future if we know its present position; it is: *All points on a wavefront can be considered as point sources for the production of spherical secondary wavelets. After a time t the new position of the wavefront will be the surface of tangency to these secondary wavelets.*

We illustrate this by a trivial example: Given a wavefront (*ab* in Fig. 41-7) in a plane wave in free space, where will the wavefront be a time *t* later? Following Huygens' principle, we let several points on this plane (see dots) serve as centers for secondary spherical wavelets. In a time *t* the radius of these spherical waves is *ct*, where *c* is the speed of light in free space. The plane of tangency to these spheres at time *t* is represented by *de*. As we

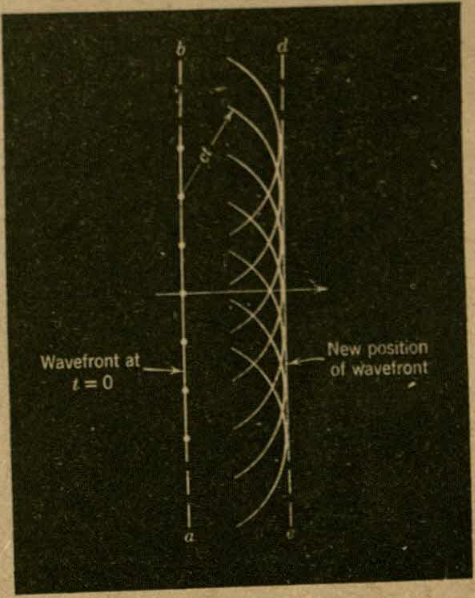


Fig. 41-7 The propagation of a plane wave in free space is described by the Huygens construction. Note that the ray (horizontal arrow) representing the wave is perpendicular to the wavefronts.

expect, it is parallel to plane *ab* and a perpendicular distance *ct* from it. Thus plane wavefronts are propagated as planes and with speed *c*. Note that the Huygens method involves a three-dimensional construction and that Fig. 41–7 is the intersection of this construction with the plane of the page.

We might expect that, contrary to observation, a wave should be radiated backward as well as forward from the dots in Fig. 41–7. This result is avoided by assuming that the *intensity* of the spherical wavelets is not uniform in all directions but varies continuously from a maximum in the forward direction to a minimum of zero in the back direction. This is suggested by the shading of the circular arcs in Fig. 41–7. Huygens' method can be applied quantitatively to *all* wave phenomena; see Problem 1. The method was put on a firm mathematical footing by Augustin Fresnel (1788–1827).

41–3 Huygens' Principle and the Law of Reflection

Figure 41–8*a* shows three wavefronts in a plane wave falling on mirror MM'. For convenience they are chosen to be one wavelength apart. Note that θ_1, the angle between the wavefronts and the mirror, is the same as the angle between the incident ray and the normal to the mirror. In other

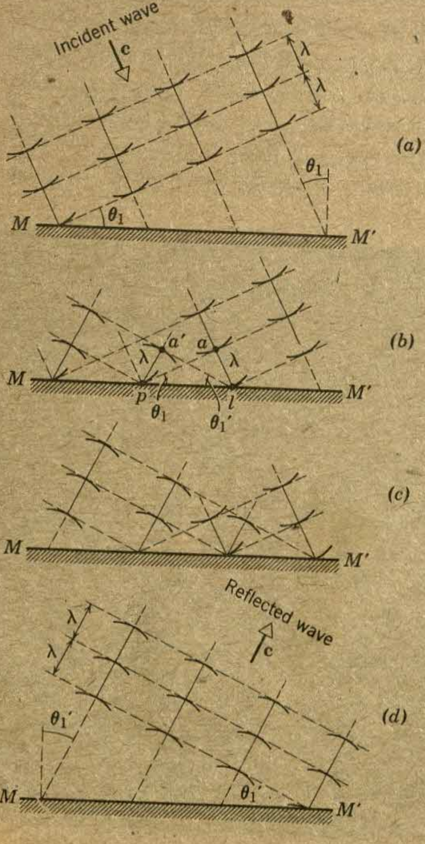


Fig. 41–8 The reflection of a plane wave as described by the Huygens construction.

words, θ_1 is the *angle of incidence*. The three wavefronts are related to each other by the Huygens construction, as in Fig. 41-7.

In Fig. 41-8*b* a Huygens wavelet centered on point *a* will expand to include point *l* after a time λ/c. Light from point *p* in this same wavefront cannot move beyond the mirror but must expand upward as a spherical Huygens wavelet. Setting a compass to radius λ and swinging an arc about *p* provides a semicircle to which the reflected wavefront must be tangent. Since point *l* must lie on the new wavefront, this tangent must pass through *l*. Note that the angle θ_1' between the wavefront and the mirror is the same as the angle between the reflected ray and the normal to the mirror. In other words, θ_1' is the *angle of reflection*.

Consider right triangles *alp* and *a'lp*. They have side *lp* in common and side *al* ($= \lambda$) is equal to side *a'p*. The two right triangles are thus congruent and we may conclude that

$$\theta_1 = \theta_1',$$

as required by the law of reflection. If the student recalls that the Huygens construction is three-dimensional and that the arcs shown represent segments of spherical surfaces, he will be able to convince himself that the reflected ray lies in the plane formed by the incident ray and the normal to the mirror, that is, the plane of Fig. 41-8. This is also a requirement of the law of reflection; see p. 1014.

41-4 Huygens' Principle and the Law of Refraction

Figure 41-9 shows four stages in the refraction of three successive wavefronts in a plane wave falling on an interface between air (medium 1) and glass (medium 2). For convenience, we assume that the incident wavefronts are separated by λ_1, the wavelength as measured in medium 1. Let the speed of light in air be v_1 and that in glass be v_2. We assume that

$$v_2 < v_1. \qquad (41\text{-}5)$$

This assumption about the speeds is vital to the derivation that follows. It was not possible to test it experimentally because of technical difficulties until 1850, when the assumption was shown by Foucault to be correct.

The wavefronts in Fig. 41-9*a* are related to each other by the Huygens construction of Fig. 41-7. As in Fig. 41-8, θ_1 is the angle of incidence. In Fig. 41-9*b* consider the time ($= \lambda_1/v_1$) during which a Huygens wavelet from point *e* moves to include point *c*. Light from point *h*, traveling through glass at a reduced speed (recall the assumption of Eq. 41-5) will move a shorter distance

$$\lambda_2 = \lambda_1 \frac{v_2}{v_1} \qquad (41\text{-}6)$$

during this time. The refracted wavefront must be tangent to an arc of this radius centered on *h*. Since *c* lies on the new wavefront, the tangent must pass through this point, as shown. Note that θ_2, the angle between the refracted wavefront and the air-glass interface, is the same as the angle

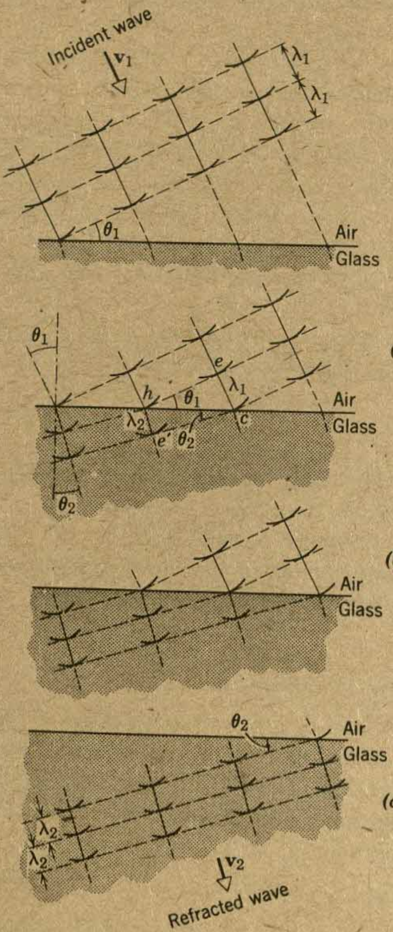


Fig. 41–9 The refraction of a plane wave as described by the Huygens construction; the reflected wave is omitted for simplicity. Note the change in wavelength on refraction.

between the refracted ray and the normal to this interface. In other words, θ_2 is the *angle of refraction*. Note, too, that the wavelength in glass (λ_2) is less than the wavelength in air (λ_1).

For the right triangles *hce* and *hce'* we may write

$$\sin \theta_1 = \frac{\lambda_1}{hc} \qquad \text{(for } hce\text{)}$$

and

$$\sin \theta_2 = \frac{\lambda_2}{hc} \qquad \text{(for } hce'\text{)}.$$

Dividing and using Eq. 41–6 yields

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{\lambda_1}{\lambda_2} = \frac{v_1}{v_2} = \text{a constant.} \tag{41–7}$$

The law of refraction, as stated in Eq. 41-2, is

$$\frac{\sin\theta_1}{\sin\theta_2} = n_{21}, \tag{41-2}$$

so that n_{21} is now revealed as the ratio of the speeds of light in the two media, or

$$n_{21} = \frac{v_1}{v_2}. \tag{41-8}$$

We may rewrite Eq. 41-7 as

$$\left(\frac{c}{v_1}\right)\sin\theta_1 = \left(\frac{c}{v_2}\right)\sin\theta_2, \tag{41-9}$$

in which c is the speed of light in free space. The quantities (c/v_1) and (c/v_2) (see Eq. 41-8) are the indices of refraction of medium 1 and of medium 2, respectively, with respect to a vacuum. Introducing the symbols n_1 and n_2 for these quantities allows us to write the law of refraction as

$$n_1 \sin\theta_1 = n_2 \sin\theta_2. \tag{41-10}$$

If we assume that the medium above the glass in Fig. 41-9 is a vacuum rather than air, the speed v_1 becomes c and the wavelength, called λ_1 in Fig. 41-9, assumes a value λ that is characteristic of the wave in free space. Equation 41-6 may thus be written

$$\lambda_2 = \lambda\frac{v_2}{c} = \frac{\lambda}{n_2}. \tag{41-11}$$

This shows specifically that the wavelength of light in a material medium is less than the wavelength of the same wave in a vacuum. Figure 41-9 shows clearly the difference in wavelength in the two media.

The application of Huygens' principle to refraction requires that if a light ray is bent toward the normal in passing from air to an optically dense medium then the speed of light in that optically dense medium (glass, say) must be *less* than that in air; see Eq. 41-5. This requirement holds for all wave theories of light. For the early particle theory of light put forward by Newton, refraction can be explained only if the speed of light in the medium in which light is bent toward the normal (the optically dense medium) is *greater* than that in air. The dense medium was thought to exert attractive forces on the light "corpuscles" as they neared the surface, speeding them up and changing their direction to cause them to make a smaller angle with the normal. Figure 41-10 shows a figure from a 1637 work of René Descartes, in which he makes an analogy between the refraction of light and the motion of a tennis ball on entering a medium in which it moves more slowly.

An experimental comparison of the speed of light in water and in air is decisive, therefore, between the wave and corpuscular theories of light. Such a measurement was first carried out by Foucault in 1850; he showed conclusively that *light travels more slowly in water than in air*, thus ruling out the corpuscular theory of Newton.

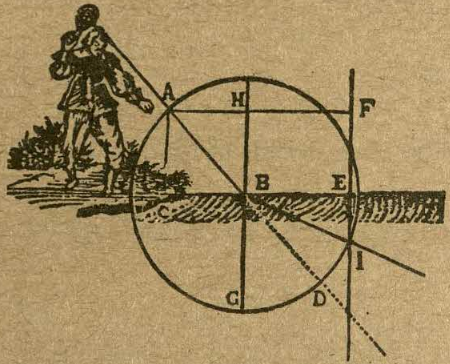


Fig. 41–10 According to the early (incorrect) corpuscular theory of light, ABI is the path of a ray on entering a medium in which its speed is *less*. The (correct) wave theory predicts that for the ray shown the speed in the lower medium (below BE) must be *greater*; from Descartes' *La Dioptrique* (1637).

41–5 Total Internal Reflection

Let light rays in an optically dense medium (glass, say) fall on a surface on the other side of which is a less optically dense medium (air, say); see Fig. 41–11. As the angle of incidence θ is increased, a situation is reached (see ray e) at which the refracted ray points along the surface, the angle of refraction being 90°. For angles of incidence larger than this *critical angle* θ_c no refracted ray exists, giving rise to a phenomenon called *total internal reflection.*

The critical angle is found by putting $\theta_2 = 90°$ in the law of refraction (see Eq. 41–10):

$$n_1 \sin \theta_c = n_2 \sin 90°,$$

or

$$\sin \theta_c = \frac{n_2}{n_1}. \qquad (41\text{–}12)$$

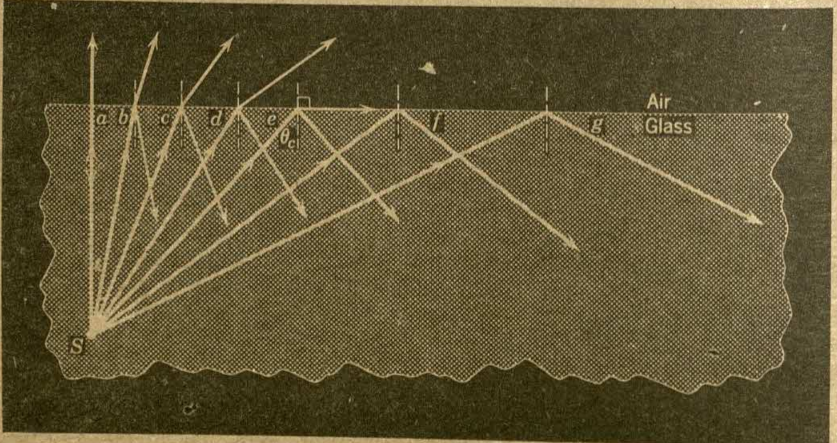


Fig. 41–11 Showing the total internal reflection of light from a source S; the *critical angle* is θ_c.

For glass and air $\sin\theta_c = (1.00/1.50) = 0.667$, which yields $\theta_c = 41.8°$. Total internal reflection does not occur when light originates in the medium of lower index of refraction.

Light can be "piped" from one point to another with little loss by allowing it to enter one end of a rod of transparent plastic. The light will undergo total internal reflection at the boundary of the rod and will follow its contour, emerging at its far end. Images may be transferred from one location to another, using a bundle of fine glass fibers, each fiber transmitting a small fraction of the image.* Such bundles can be made in which a (flexible) seven-foot length delivers half the energy entering it at the far end. In a fiber in such a bundle a typical ray may undergo 48,000 reflections. Most of the energy loss is due to absorption within the glass, reflection being almost

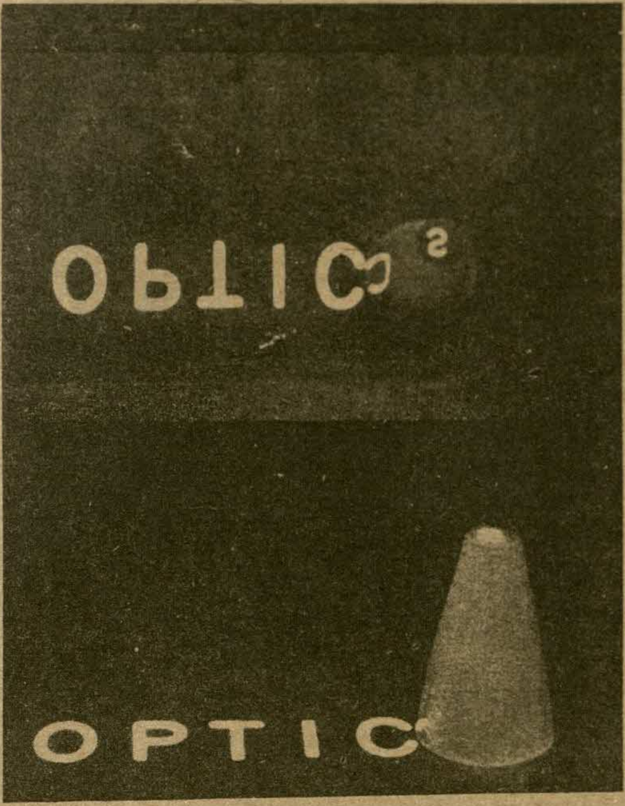


Fig. 41–12 A bundle of tapered fibers (*below*) is placed over the letter *S*. Above, with the aid of a mirror, we see that the image, reduced in size, is transmitted to the top of the bundle by total internal reflection in the individual fibers. (Courtesy Dr. N. S. **Kapany**, Optics Technology, Inc.)

---

* See "Fiber Optics," by N. S. Kapany, *Scientific American*, November 1960.

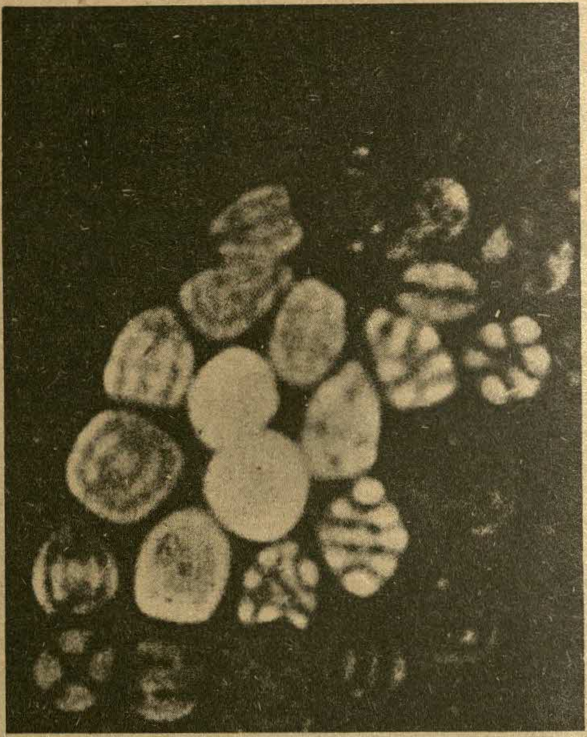


Fig. 41–13 A photomicrograph of light emerging from the end of a bundle of fibers. The fiber diameters approach the wavelength of light so that each fiber acts like an "optical waveguide." We have here convincing visual evidence that light is an electromagnetic wave. (Courtesy of Dr. N. S. Kapany, Optics Technology, Inc.)

truly total. Measurable amounts of light can be transmitted through single fibers 150 feet long.

Fiber optics techniques make possible many useful optical devices for transmitting and transforming luminous images. Figure 41–12 shows a short fiber bundle constructed so that the fibers taper in diameter along its length. The wide end is shown placed over the letter *S* in the printed word "OPTICS." We see, with the aid of a mirror placed above the bundle, that a letter *S*, reduced in size, has been transmitted by total internal reflection through the bundle to its narrow end.

Figure 41–13 is an enlarged view of a cross section of such a fiber bundle in which the diameters of the individual fibers are made so small that they are of the order of magnitude of the wavelength of light. This condition violates the spirit of our assumption of p. 1016, namely that the transverse dimensions of reflecting and refracting surfaces would be large compared to the wavelength of light. Consequently, a description of the reflection and refraction of light in terms of rays, as in Figs. 41–1*b* and 41–11, is not possible. Figure 41–13 is readily interpreted, however, on the basis of the electromagnetic wave theory of light and provides convincing pictorial supporting evidence for that theory. The fibers behave like waveguides *

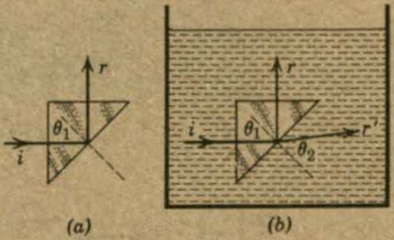
---

* A dielectric rod can serve as a waveguide in the same way that a hollow metal pipe can, and it has similar properties.

(see Section 39-3), and the patterns of darkness and light represent the distribution of the **E** and **B** vectors for various modes of oscillation of the electromagnetic waves traveling down the "guides."

▶ **Example 4.** Figure 41-14*a* shows a triangular prism of glass, a ray incident normal to one face being totally reflected. If θ_1 is 45°, what can you conclude about the index of refraction n of the glass?

Fig. 41-14 Example 4.



The angle θ_1 must be equal to or greater than the critical angle θ_c where θ_c is given by Eq. 41-12:

$$\sin\theta_c = \frac{n_2}{n_1} = \frac{1}{n},$$

in which, for all practical purposes, the index of refraction of air ($= n_2$) is set equal to unity. Suppose that the index of refraction of the glass is such that total internal reflection just occurs, that is, that $\theta_c = 45°$. This would mean

$$n = \frac{1}{\sin 45°} = 1.41.$$

Thus the index of refraction of the glass must be *equal to or larger than* 1.41. If it were less, total internal reflection would not occur.

Example 5. What happens if the prism in Example 4 (assume that $n = 1.50$) is immersed in water ($n = 1.33$)? See Fig. 41-14*b*.

The new critical angle, given by Eq. 41-12, is

$$\sin\theta_c = \frac{n_2}{n_1} = \frac{1.33}{1.50} = 0.887,$$

which corresponds to $\theta_c = 62.5°$. The actual angle of incidence ($= 45°$) is less than this so that we do *not* have total internal reflection.

There is a reflected ray, with an angle of reflection of 45°, as Fig. 41-14*b* shows. There is also a refracted ray, with an angle of refraction given by

$$n_1 \sin\theta_1 = n_2 \sin\theta_2$$

$$(1.50)(\sin 45°) = (1.33)\sin\theta_2,$$

which yields $\theta_2 = 52.9°$. ◀

Maxwell's equations permit us to calculate how the incident energy is divided between the reflected and the refracted beams. Figure 41-15 shows the theoretical

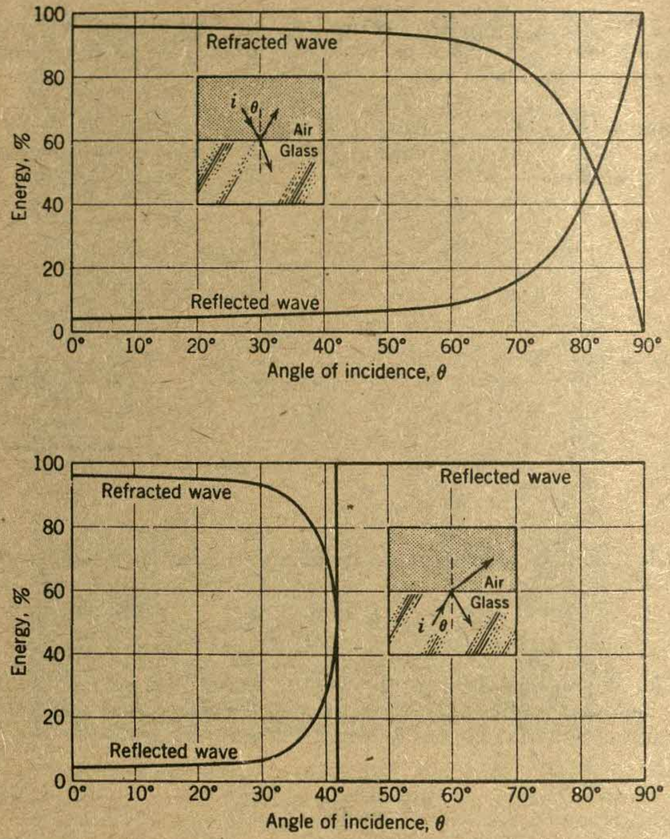


Fig. 41–15 (*a*) The per cent of the energy reflected and refracted when an incident wave in air falls on glass ($n = 1.50$). (*b*) The same for the incident wave in glass, showing total internal reflection.

prediction for (*a*) a light beam in air falling on a glass-air interface and (*b*) a light beam in glass falling on such an interface. Figure 41–15*a* shows that for angles of incidence up to about 50°, less than 10% of the light energy is reflected. At grazing incidence, however (that is, angles of incidence near 90°), the surface becomes an excellent reflector. We are all familiar with the high reflecting power of a wet road for light from automobile headlights that strikes near grazing incidence.

Figure 41–15*b* shows clearly that at a certain critical angle (41.8° in this case; see Eq. 41–12) *all* the light is reflected. For angles of incidence appreciably below this value, about 4% of the energy is reflected.

41–6 Fermat's Principle

In 1650 Pierre Fermat discovered a remarkable principle which we often express today in these terms: *A light ray traveling from one point to another will follow a path such that, compared with nearby paths, the time required is either a minimum or a maximum or will remain unchanged (that is, it will be stationary).*

The laws of reflection and refraction can readily be derived from this principle.

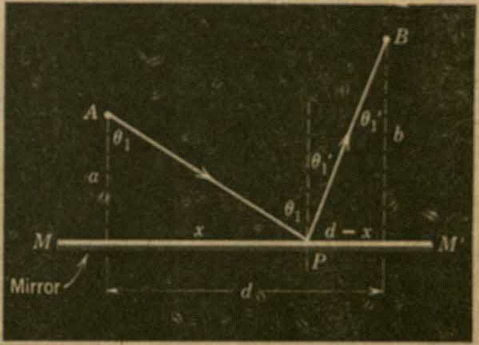


Fig. 41–16 A ray from A passes through B after reflection at P.

Figure 41–16 shows two fixed points A and B and a ray APB connecting them.* The total length l of this ray is

$$l = \sqrt{a^2 + x^2} + \sqrt{b^2 + (d - x)^2},$$

where x locates the point P at which the ray touches the mirror.

According to Fermat's principle, P will have a position such that the time of travel of the light must be a minimum (or a maximum or must remain unchanged). Expressed in another way, the total length l of the ray must be a minimum (or a maximum or must remain unchanged). In either case, the methods of the calculus require that dl/dx be zero. Taking this derivative yields

$$\frac{dl}{dx} = (\tfrac{1}{2})(a^2 + x^2)^{-1/2}(2x) + \tfrac{1}{2}[b^2 + (d - x)^2]^{-1/2}(2)(d - x)(-1) = 0$$

which can be rewritten as

$$\frac{x}{\sqrt{a^2 + x^2}} = \frac{d - x}{\sqrt{b^2 + (d - x)^2}}.$$

Comparison with Fig. 41–16 shows that this can be written as

$$\sin \theta_1 = \sin \theta_1',$$

or

$$\theta_1 = \theta_1',$$

which is the law of reflection.

To prove the law of refraction from Fermat's principle, consider Fig. 41–17, which shows two points A and B in two different media and a ray APB connecting them. The time t is given by

$$t = \frac{l_1}{v_1} + \frac{l_2}{v_2}.$$

Using the relation $n = c/v$ this can be written as

$$t = \frac{n_1 l_1 + n_2 l_2}{c} = \frac{l}{c}.$$

The quantity l ($= n_1 l_1 + n_2 l_2$) is called the *optical path length* of the ray. Equation 41–11 (written as $\lambda = n\lambda_n$) shows that the optical path length is equal to the length

---

* We assume that ray APB lies in the plane of the figure; see Problem 20.

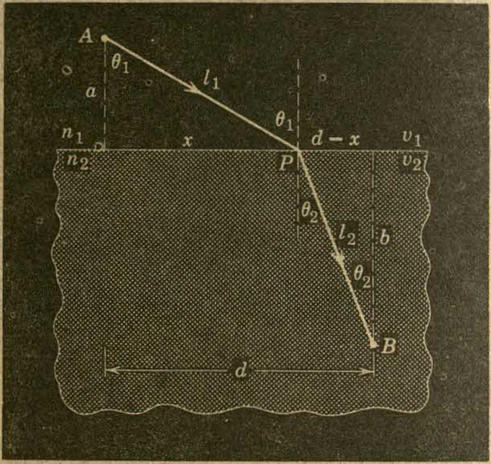


Fig. 41–17 A ray from A passes through B after refraction at P.

that this same number of waves would have if the medium were a vacuum. The optical path length must not be confused with the geometrical path length, which is $l_1 + l_2$.

Fermat's principle requires that l be a minimum (or a maximum or must remain unchanged) which, in turn, requires that x be so chosen that $dl/dx = 0$. The optical path length is

$$l = n_1 l_1 + n_2 l_2 = n_1\sqrt{a^2 + x^2} + n_2\sqrt{b^2 + (d - x)^2}.$$

Differentiating yields

$$\frac{dl}{dx} = n_1(\tfrac{1}{2})(a^2 + x^2)^{-1/2}(2x) + n_2(\tfrac{1}{2})[b^2 + (d - x)^2]^{-1/2}(2)(d - x)(-1) = 0,$$

which can be rewritten as

$$n_1 \frac{x}{\sqrt{a^2 + x^2}} = n_2 \frac{d - x}{\sqrt{b^2 + (d - x)^2}}.$$

Comparison with Fig. 41–17 shows that this, in turn, can be written as

$$n_1 \sin \theta_1 = n_2 \sin \theta_2,$$

which is the law of refraction.

In each of the examples of this section the time required, or, what is equivalent, the optical path length, proves to be a *minimum*. Problem 22 describes a case in which it may be a maximum, a minimum, or stationary.

QUESTIONS

1. Discuss the propagation of spherical and of cylindrical waves, using Huygens' principle.

2. Does Huygens' principle apply to sound waves in air?

3. If Huygens' principle predicts the laws of reflection and refraction, why is it necessary or desirable to view light as an electromagnetic wave, with all its attendant complexity?

4. Would you expect sound waves to obey the laws of reflection and of refraction obeyed by light waves?

5. A street light, viewed by reflection across a body of water in which there are ripples, appears very elongated. Explain.

6. The light beam in Fig. 41–1*a* is broadened on entering the water. Explain.

7. By what per cent does the speed of blue light in fused quartz differ from that of red light?

8. Can (*a*) reflection phenomena or (*b*) refraction phenomena be used to determine the wavelength of light?

9. How did Foucault's measurement of the speed of light in water decide between the wave and the particle theories of light?

10. Why does a diamond "sparkle" more than a glass imitation cut to the same shape?

11. Is it plausible that the wavelength of light should change in passing from air into glass but that its frequency should not? Explain.

12. How can one determine the indices of refraction of the media in Table 41–1 relative to water, given the data in that table?

13. You are given a cube of glass. How can you find the speed of light (from a sodium light source) in this cube?

14. Describe and explain what a fish sees as he looks in various directions above his "horizon."

15. Design a periscope, taking advantage of total internal reflection. What are the advantages compared with silvered mirrors?

16. What characteristics must a material have in order to serve as an efficient "light pipe"?

17. Can the optical path length between two points ever be less than the geometrical path length between those points?

PROBLEMS

1. One end of a stick is dragged through water at a speed v which is greater than the speed u of water waves. Applying Huygens' construction to the water waves, show that a conical wavefront is set up and that its half-angle α is given by

$$\sin \alpha = u/v.$$

This is familiar as the bow wave of a ship or the shock wave caused by an object moving through air with a speed exceeding that of sound, as in Fig. 20–12.

2. Prove that if a mirror is rotated through an angle α, the reflected beam is rotated through an angle 2α. Is this result reasonable for $\alpha = 45°$?

3. What is the speed in fused quartz of light of wavelength 5500 A; see Fig. 41–2?

4. The wavelength of yellow sodium light in air is 5890 A. (*a*) What is its frequency? (*b*) What is its wavelength in glass whose index of refraction is 1.52? (*c*) From the results of (*a*) and (*b*) find its speed in this glass.

5. The speed of yellow sodium light in a certain liquid is measured to be 1.92×10^8 meters/sec. What is the index of refraction of this liquid, with respect to air, for sodium light?

6. Suppose that the speed of light in air has been measured with an uncertainty of, say, 1 km/sec. In calculating the speed in vacuum, suppose that it is not certain whether n for air is 1.00029 or 1.00030. (*a*) How much extra uncertainty is introduced into the calculated value for c? (*b*) Estimate how accuratel[illegible]ould be known for this purpose.

7. In Fig. 41–6 show by graphical ray tracing, using a protractor, that if θ for the incident ray is *either* increased *or* decreased, the deviation angle ψ is increased. The symmetrical situation shown in this figure is called the *position of minimum deviation.*

8. A ray of light is incident normally on the face *ab* of a glass prism ($n = 1.52$) as shown in Fig. 41-18. (*a*) Assuming that the prism is immersed in air, find the largest value for the angle ϕ so that the ray is totally reflected at face *ac*. (*b*) Find ϕ if the prism is immersed in water.

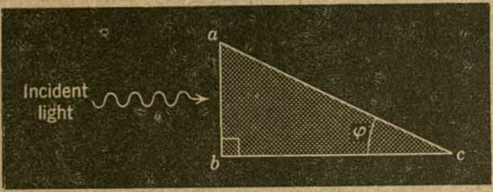


Fig. 41-18

9. Show that for a thin prism (ϕ small) and light not far from normal incidence (θ_1 small) the deviation angle is independent of the angle of incidence and is equal to $(n - 1)\,\phi$ (see Fig. 41-6).

10. A glass prism with an apex angle of 60° has $n = 1.60$. (*a*) What is the smallest angle of incidence for which a ray can enter one face of the prism and emerge from the other? (*b*) What angle of incidence would be required for the ray to pass through the prism symmetrically, as in Fig. 41-6?

11. A 60° prism is made of fused quartz. A ray of light falls on one face, making an angle of 45° with the normal. Trace the ray through the prism graphically with some care, showing the paths traversed by rays representing (*a*) blue light, (*b*) yellow-green light, and (*c*) red light. See Figs. 40-2 and 41-2.

12. Ptolemy, who lived at Alexandria toward the end of the first century A.D., gave the following measured values for the angle of incidence θ_1 and the angle of refraction θ_2 for a light beam passing from air to water:

θ_1	θ_2	θ_1	θ_2
10°	7°45′	50°	35°0′
20°	15°30′	60°	40°30′
30°	22°30′	70°	45°30′
40°	29°0′	80°	50°0′

Are these data consistent with Snell's law; if so, what index of refraction results? These data are interesting as the oldest recorded physical measurements.

13. Prove that a ray of light incident on the surface of a sheet of plate glass of thickness t emerges from the opposite face parallel to its initial direction but displaced sideways, as in Fig. 41-19. Show that, for small angles of incidence θ, this displacement is given by

$$x = t\theta\frac{n - 1}{n}$$

where n is the index of refraction and θ is measured in radians.

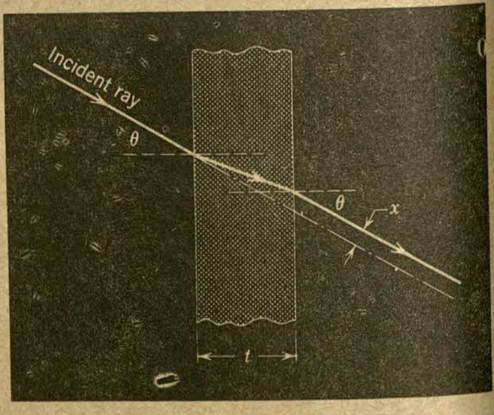


Fig. 41-19

14. A plane wave of white light traveling in fused quartz strikes a plane surface of the quartz, making an angle of incidence θ. Is it possible for the internally reflected beam to appear (*a*) bluish or (*b*) reddish? Roughly what value of θ must be used? (Hint: White light will appear bluish if wavelengths corresponding to red are removed from the spectrum.)

15. A point source of light is placed a distance h below the surface of a large deep lake. (*a*) Show that the fraction f of the light energy that escapes directly from the water surface

is independent of h and is given by

$$f = \frac{1}{2} - \frac{1}{2n}\sqrt{n^2 - 1}$$

where n is the index of refraction of water. (Note: Absorption within the water and reflection at the surface—except where it is total—have been neglected.) (*b*) Evaluate this fraction for $n = 1.33$.

16. Figure 41-20 shows a *constant-deviation prism*. Although made of one piece of glass, it is equivalent to two 30°-60°-90° prisms and one 45°-45°-90° prism. White light is incident in the direction i. θ_1 is changed by rotating the prism so that, in turn, light of any desired wavelength may be made to follow the path shown, emerging at r. Show that, if $\sin\theta_1 = \frac{1}{2}n$, then $\theta_2 = \theta_1$ and beams i and r are at right angles.

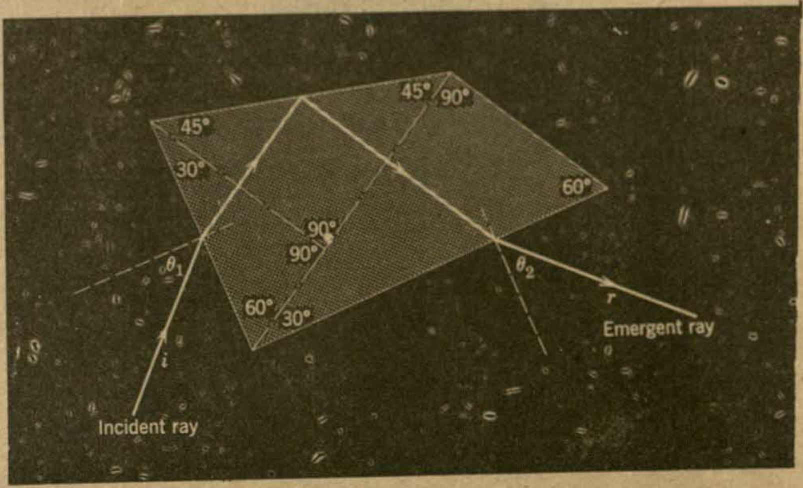


Fig. 41-20

17. A drop of liquid may be placed on a semicircular slab of glass as in Fig. 41-21. Show how to determine the index of refraction of the liquid by observing total internal reflection. The index of refraction of the glass is unknown and must also be determined. Is the range of indices of refraction that can be measured in this way restricted in any sense?

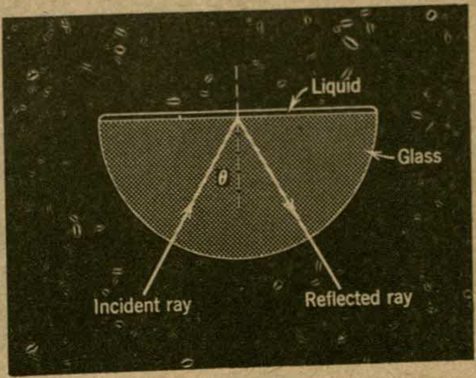


Fig. 41-21

18. A light ray falls on a square glass slab as in Fig. 41–22. What must the index of refraction of the glass be if total internal reflection occurs at the vertical face?

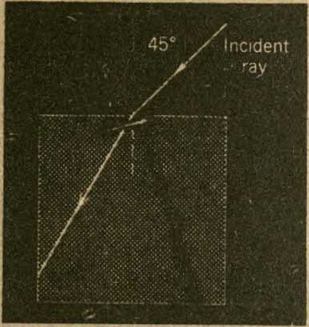


Fig. 41–22

19. A glass cube has a small spot at its center. What parts of the cube face must be covered to prevent the spot from being seen, no matter what the direction of viewing? What fraction of the cube surface must be so covered? Assume a cube edge of 1.0 cm and an index of refraction of 1.50. (Neglect the subsequent behavior of an internally reflected ray.)

20. Using Fermat's principle, prove that the reflected ray, the incident ray, and the normal lie in one plane.

21. Prove that the optical path lengths for reflection and refraction in Figs. 41–16 and 41–17 are minima when compared with other nearby paths connecting the same two points.

22. Figure 41–23 shows two points A and B connected by a light ray AvB. Show that, by comparison with nearby rays, such as AcB, the ray AvB represents a minimum, a stationary, or a maximum optical path, depending on whether the distance a is respectively less than, equal to, or greater than, the quantity

$$R\frac{n+1}{n-1}.$$

R is the radius of curvature of the spherical surface and n is the index of refraction of the medium to the right of this surface. (Hint: Express the optical path length l of the arbitrary path AcB in terms of the angle ϕ in Fig. 41–23. Use approximations that are appropriate when ϕ is small, since AcB must be a *nearby* path. Take the first and second derivatives of l with respect to ϕ.)

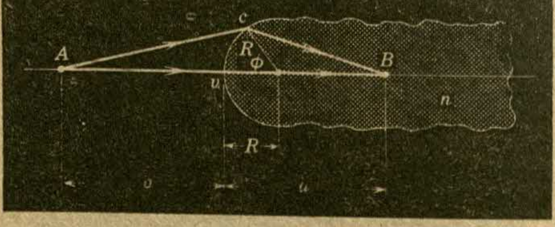


Fig. 41–23

Reflection and Refraction—Spherical Waves and Spherical Surfaces

CHAPTER 42

42-1 Geometrical Optics and Wave Optics

In Chapter 41 we described the reflection and refraction of plane waves at plane surfaces. In this chapter we consider the more general case of spherical waves falling on spherical reflecting and refracting surfaces. All of the results of Chapter 41 will emerge as special cases of the results of this chapter, since a plane can be viewed as a spherical surface with an infinite radius of curvature.

Both in Chapter 41 and in this chapter we make extensive use of *rays*. Although a ray is a convenient construction, it proves impossible to isolate one physically. Figure 42–1a shows schematically a plane wave of wavelength λ falling on a slit of width $a = 5\lambda$. We find that the light flares out into the geometrical shadow of the slit, a phenomenon called *diffraction*. Figures 42–1b ($a = 3\lambda$) and 42–1c ($a = \lambda$) show that diffraction becomes more pronounced as $a/\lambda \rightarrow 0$ and that attempts to isolate a single ray from the incident plane wave are futile.

Figure 42–2 shows water waves in a shallow *ripple tank*, produced by tapping the water surface periodically and automatically with the edge of a flat stick. We see that the plane wave so generated flares out by diffraction when it encounters a gap in a barrier placed across it. Diffraction is characteristic of waves of all types. We can hear around corners, for example, because of the diffraction of sound waves.

The diffraction of waves at a slit (or at an obstacle such as a wire) is expected from Huygens' principle. Consider the portion of the wavefront that arrives at the position of the slit in Fig. 42–1. Every point on it can be

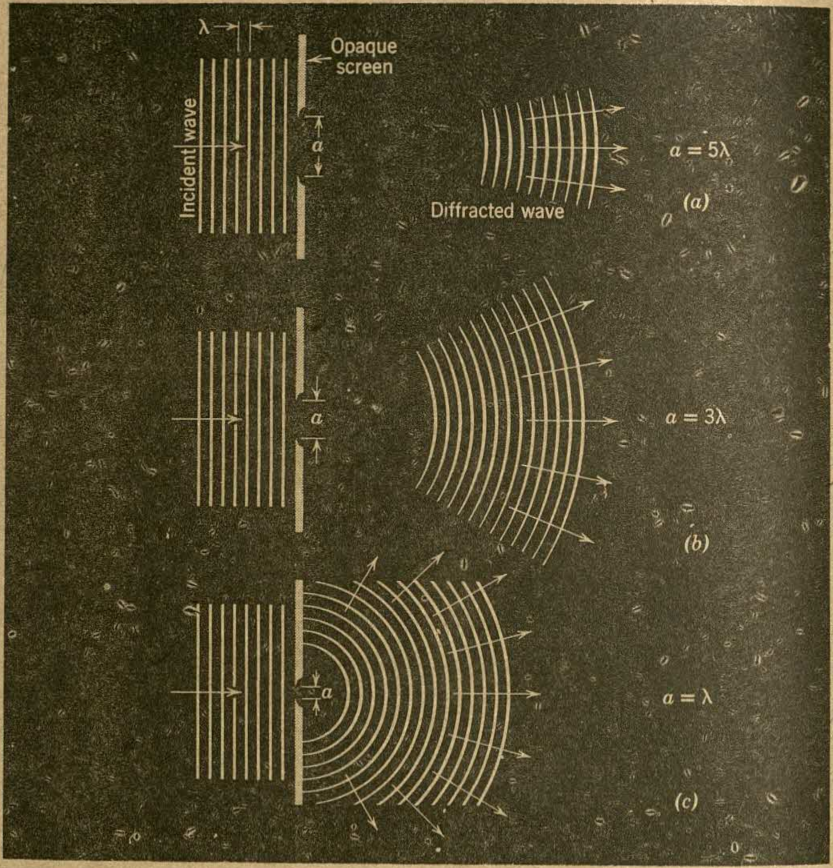


Fig. 42–1 An attempt to isolate a ray by reducing the slit width a fails because of diffraction, which becomes more pronounced as a/λ approaches zero.

viewed as the site of an expanding spherical Huygens' wavelet. The "bending" of light into the region of the geometrical shadow is associated with the blocking off of Huygens' wavelets from those parts of the incident wavefront that lie behind the slit edges.

Figure 42–3 was made by allowing parallel light to fall on a slit placed 50 cm in front of a photographic plate. In Fig. 42–3a the slit width was about 6×10^{-3} mm. The central band of light is much wider than this, showing that light has "flared out" into the geometric shadow of the slit. In addition, many secondary maxima, omitted from Fig. 42–1 for simplicity, appear. Figure 42–3b shows what happens when the slit width is *reduced* by a factor of two. The central maximum becomes *wider*, in agreement with Fig. 42–1. Figure 42–3c shows the effect of reducing the slit width by an additional factor of 7, to 4×10^{-4} mm. The central maximum is now much wider and the secondary maxima, whose intensities relative to the central

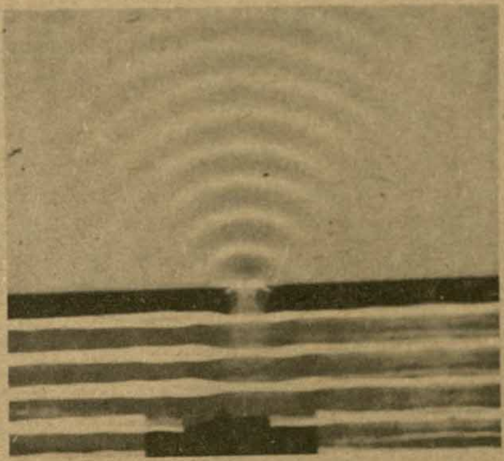


Fig. 42-2 Diffraction of water waves at a slit in a ripple tank. Note that the slit width is about the same size as the wavelength. (Courtesy of Educational Services Incorporated.)

maximum have been deliberately overemphasized by long exposure, are very evident.

Diffraction can be ignored if the ratio a/λ is large enough, a being a measure of the smallest sideways dimension of the slit or obstacle. If $a \gg \lambda$, light appears to travel in straight lines which can be represented by rays that obey the laws of reflection (Eq. 41-1) and refraction (Eq. 41-2). In Chapter 41 this condition, called *geometrical optics*, prevailed, the lateral dimensions of all mirrors, prisms, etc., being much greater than the wavelength. We assume in this chapter also that the conditions for geometrical optics are satisfied.

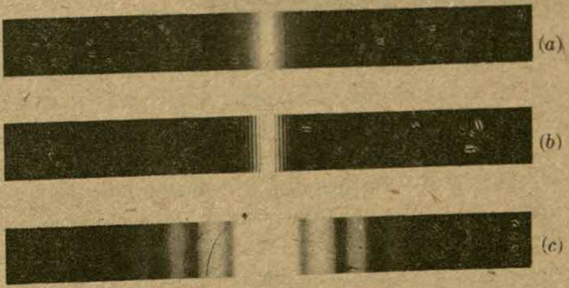


Fig. 42-3 (*a*) The intensity of light diffracted from a slit of width $a \cong 6 \times 10^{-3}$ mm and falling on a screen 50 cm beyond. (*b*) The slit width is reduced by a factor of two. (*c*) The slit width is further reduced by an additional factor of seven. Note that secondary maxima, made prominent in this case by deliberate overexposure, appear on either side of the central maximum. These secondary maxima have been omitted from Fig. 42-1 for simplicity.

If the requirement for geometrical optics is not met, we cannot describe the behavior of light by rays but must take its wave nature specifically into account. This subject is called *wave optics;* it includes geometrical optics as an important limiting case. We will treat wave optics in succeeding chapters.

42–2 Spherical Waves—Plane Mirror

Figure 42–4 shows a point source of light *O*, the *object*, placed a distance *o* in front of a plane mirror. The light falls on the mirror as a spherical wave represented in the figure by rays emanating from *O*.* At the point at which each ray strikes the mirror we construct a reflected ray. If the reflected rays are extended backward, they intersect in a point *I* which is the same distance behind the mirror that the object *O* is in front of it; *I* is called the *image* of *O*.

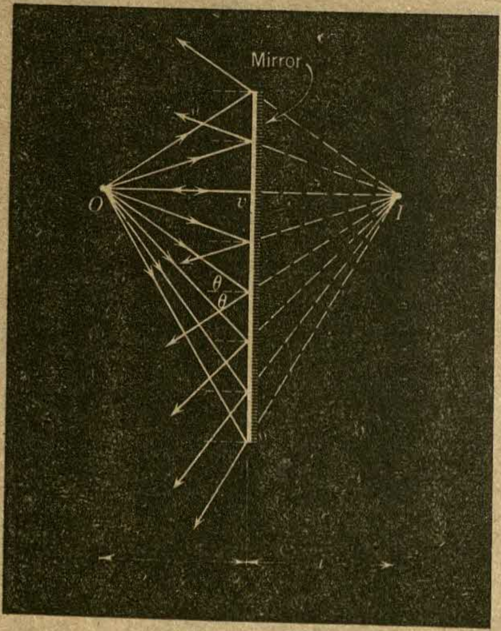


Fig. 42–4 A point object *O* forms a virtual image *I* in a plane mirror. The rays appear to emanate from *I*, but actually light energy does not pass through this point.

Images may be *real* or *virtual*. In a real image light energy actually passes through the image point; in a virtual image the light behaves as though it diverges from the image point, although, in fact, it does not pass through this point; see Fig. 42–4. Images in plane mirrors are always virtual. We know from daily experience how "real" such a virtual image appears to be and how definite is its location in the space behind the mirror, even though this space may, in fact, be occupied by a brick wall.

Figure 42–5 shows two rays from Fig. 42–4. One strikes the mirror at *v*, along a perpendicular line. The other strikes it at an arbitrary point *a*, making an angle of incidence θ with the normal at that point. Elementary

---

* In our discussion of reflection from mirrors in Chapter 41 (see Fig. 41–8) we assumed an incident *plane* wave; the incident rays are parallel to each other in that case.

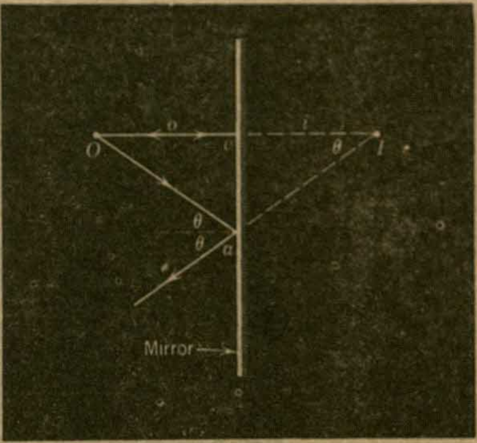


Fig. 42-5 Two rays from Fig. 42-4; ray Oa makes an arbitrary angle θ with the normal.

geometry shows that the angles aOv and aIv are also equal to θ. Thus the right triangles $aOva$ and $aIva$ are congruent and

$$o = -i, \tag{42-1}$$

in which the minus sign is arbitrarily introduced to show that I and O are on opposite sides of the mirror. Equation 42-1 does not involve θ, which means that *all* rays striking the mirror pass through I when extended backward, as we have seen above. Beyond assuming that the mirror is truly plane and that the conditions for geometrical optics hold, we have made no approximations in deriving Eq. 42-1. A point object produces a point image in a plane mirror, with $o = -i$, no matter how large the angle θ in Fig. 42-5.

Because of the finite diameter of the pupil of the eye, only rays that lie fairly close together can enter the eye after reflection at a mirror. For the eye position shown in Fig. 42-6 only a small patch of the mirror near point a is effective in forming the image; the rest of the mirror may be covered up or removed. If the eye is moved to another location, a different patch of the mirror will be effective; the location of the virtual image I will remain unchanged, however, as long as the object remains fixed.

If the object is an extended source such as the head of a person, a virtual image is also formed. From Eq. 42-1, every point of the source has an image point that lies an equal distance directly behind the plane of the mirror. Thus the image reproduces the object point by point.

Images in plane mirrors differ from objects in that left and right are interchanged. The image of a printed page is different from the page itself. Similarly, if a top is made to spin clockwise, the image, viewed in a vertical mirror, will seem to spin counterclockwise. Figure 42-7 shows an image of a left hand, constructed by using point-by-point application of Eq. 42-1; the image has the symmetry of a right hand.*

---

* See "The Overthrow of Parity" by Philip Morrison, *Scientific American*, April, 1957, for a discussion of the distinction in nature between right and left.

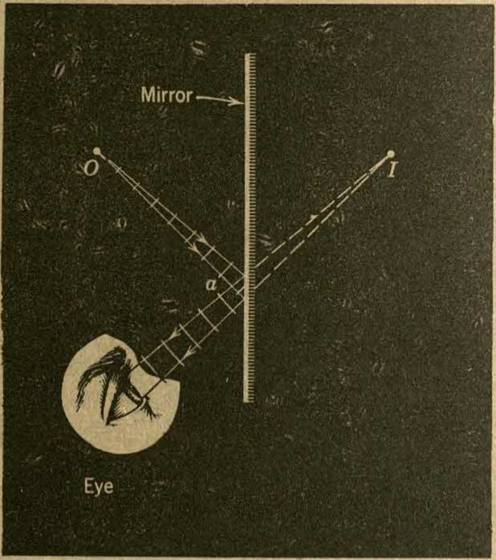


Fig. 42–6 A pencil of rays from O enters the eye after reflection at the mirror. Only a small portion of the mirror near a is effective. The small arcs represent portions of spherical wavefronts.

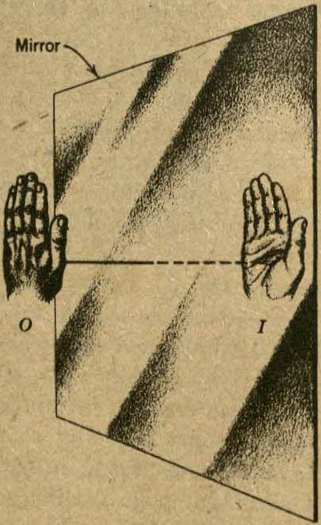


Fig. 42–7 A plane mirror reverses right and left.

▶ **Example 1.** How tall must a vertical mirror be if a person 6 ft high is to be able to see his entire length? Assume that his eyes are 4 in. below the top of his head.

Figure 42-8 shows the paths followed by light rays leaving the top of the man's head and the tips of his toes. These rays, chosen so that they will enter the eye *e* after reflection, strike the vertical mirror at points *a* and *b*, respectively. The mirror need occupy only the region between these two points. Calculation shows that *b* is 2 ft, 10 in. and *a* is 5 ft, 10 in. above the floor. The length of the mirror is thus 3 ft, or half the height of the person. Note that this height is independent of the distance between the person and the mirror. Mirrors that extend below point *b* show reflections of the floor between the person and the mirror.

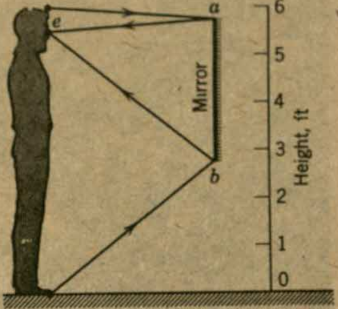


Fig. 42-8 Example 1. A person can view his full-length image in a mirror that is only half his height.

Example 2. Two plane mirrors are placed at right angles, and a point object O is located on the perpendicular bisector, as shown in Fig. 42-9*a*. Locate the images. Images I_1 and I_2 are formed in mirrors *ab* and *cd*, respectively. There is also a third image; it may be considered to be the image of I_1 in mirror *cd* or the image of I_2 in mirror *ab*. The three images and the object O lie on a circle whose center is on the line of intersection of the mirrors and whose plane is at right angles to that line.

In viewing I_3 the light entering the observer's eye is reflected *twice* after leaving the source. Figure 42-9*b* shows a typical bundle of rays. In viewing I_1 or I_2, the light is reflected only once, as in Fig. 42-6. ◀

42-3 Spherical Waves—Spherical Mirror

In Fig. 42-10 a spherical light wave from a point object O falls on a concave spherical mirror whose radius of curvature is r.* A line through O and the center of curvature C makes a convenient reference axis.

A ray from O that makes an arbitrary angle α with this axis intersects the axis at I after reflection from the mirror at a. A ray that leaves O along the axis will be reflected back along itself at v and will also pass through I. Thus, for these two rays at least, I is the image of O; it is a *real* image because light energy actually passes through I. Let us find the location of I.

A useful theorem is that the exterior angle of a triangle is equal to the sum of the two opposite interior angles. Applying this to triangles *OaCO* and

---

* A spherical shell, viewed from inside, is everywhere *concave;* viewed from outside it is everywhere *convex.* In this chapter concave and convex will always be judged from the point of view of an observer sighting along the direction of the incident light.

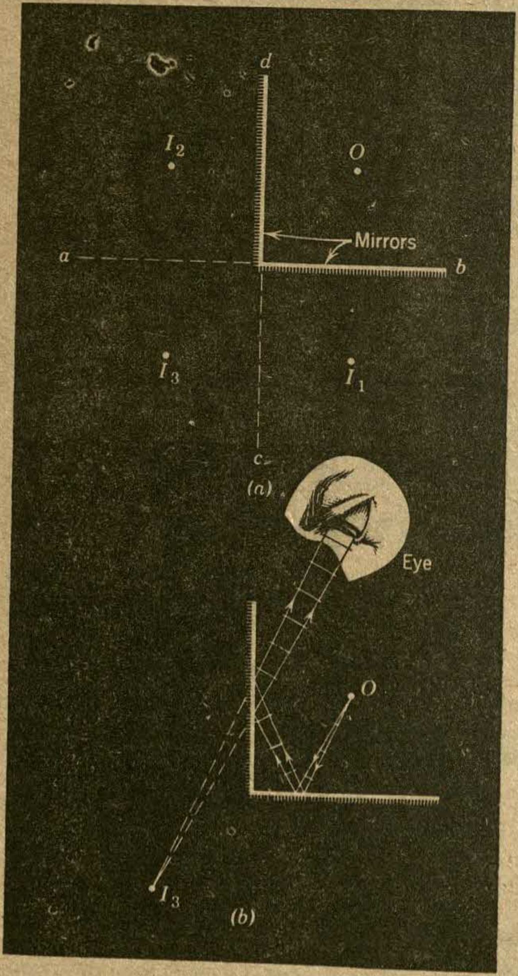


Fig. 42–9 Example 2. (*a*) Object O has three virtual images. (*b*) A typical bundle of rays used to view I_3.

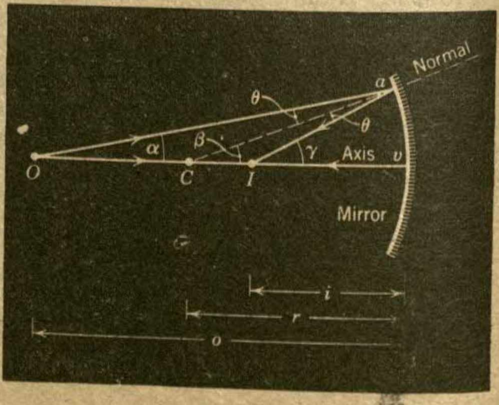


Fig. 42–10 Two rays from O converge after reflection in a spherical concave mirror, forming a real image at I.

$OaIO$ in Fig. 42–10 yields

$$\beta = \alpha + \theta$$

and

$$\gamma = \alpha + 2\theta.$$

Eliminating θ between these equations leads to

$$\alpha + \gamma = 2\beta. \tag{42-2}$$

In radian measure we can write angles α, β, and γ as

$$\begin{aligned} \alpha &\cong \frac{av}{vO} = \frac{av}{o} \\ \beta &= \frac{av}{vC} = \frac{av}{r} \\ \gamma &\cong \frac{av}{vI} = \frac{av}{i}. \end{aligned} \tag{42-3}$$

Note that only the equation for β is exact, for the reason that the center of curvature of arc av is at C and not at O or I. However, the equations for α and for γ are approximately correct if these angles are sufficiently small. *In all that follows we assume that the rays diverging from the object make only a small angle α with the axis of the mirror.* Such rays, which lie close to the mirror axis, are called *paraxial rays*. We did not find it necessary to make such an assumption for plane mirrors. Substituting these equations into Eq. 42–2 and canceling av yields

$$\frac{1}{o} + \frac{1}{i} = \frac{2}{r}, \tag{42-4}$$

in which o is the *object distance* and i is the *image distance*. Both these distances are measured from the *vertex* of the mirror, which is the point v at which the axis intercepts the mirror.

Significantly, Eq. 42–4 does not contain α (or β, γ, or θ), so that it holds for all rays that strike the mirror provided that they are sufficiently paraxial. In an actual case the rays can be made as paraxial as one likes by putting a circular diaphragm in front of the mirror, centered about the vertex v; this will impose a certain maximum value of α.

As α in Fig. 42–10 is permitted to become larger, it will become less true that a point object will form a point image; the image will become extended and fuzzy. No sharp criterion for deciding whether a given ray is paraxial can be laid down. If the maximum permitted value of α is reduced, the rays will become more paraxial and the image will become sharper. Unfortunately, the image will also become fainter because less total light energy will be reflected from the mirror. A compromise must often be made between image brightness and image quality.

As for plane mirrors, the image (real or virtual) in a spherical mirror can be seen only if the eye is located so that light rays from the object can enter

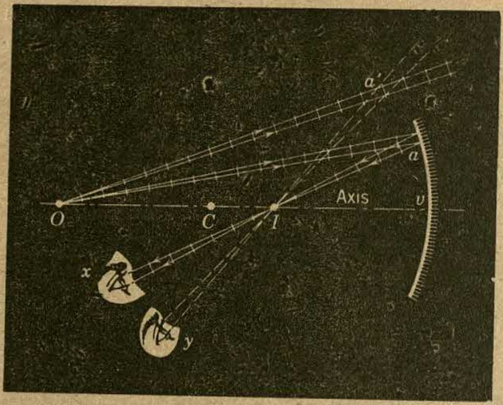


Fig. 42–11 The eye must be properly located to see image I.

it after reflection. In Fig. 42–11 a bundle of light rays is shown entering the eye in position x; only the small patch of mirror near a is effective for this eye position. If the observer moves his eye to position y, the image will vanish for him because the mirror does not exist near point a'.

Although Eq. 42–4 was derived for the special case in which the object is located beyond the center of curvature, it is generally true, no matter where the object is located. It is also true for convex mirrors, as in Fig. 42–12.

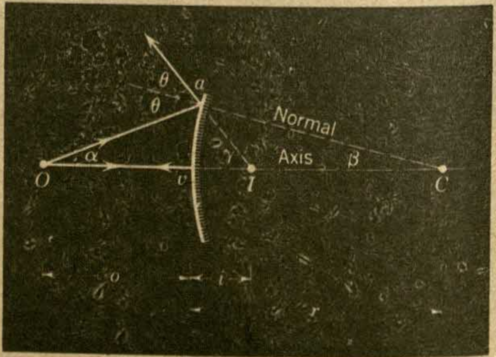
In applying Eq. 42–4, we must be careful to follow a consistent convention of signs for o, i, and r. As the basis for the sign conventions to be used in this book, we start from this statement:

> In Fig. 42–10, in which light *diverges* from a *real* object, falls on a *concave* mirror, and *converges* after reflection to form a *real* image, the quantities o, i, and r in Eq. 42–4 are given positive numerical values.

Figure 42–10 was used to derive Eq. 42–4 and the student should associate them in his mind as an aid in getting the signs correct.

Let us fix our minds on the side of the mirror from which incident light comes. Because mirrors are opaque, the light, after reflection, must remain on this side, and if an image is formed here it will be a *real* image. There-

Fig. 42–12 Two rays from O diverge after reflection in a spherical convex mirror, forming a virtual image at I, the point from which they appear to originate. Compare Fig. 42–10.



fore, we call the side of the mirror from which the light comes the *R-side* (for *real image*). We call the back of the mirror the *V-side* (for *virtual image*), because images formed on this side of the mirror must be virtual, no light energy being present on this side.*

In the indented statement above we associated real images with positive image distances. This suggests our first sign convention:

1. The image distance i is positive if the image (real) lies on the R-side of the mirror, as in Fig. 42–10; i is negative if the image (virtual) lies on the V-side, as in Fig. 42–12.

If the mirror in Fig. 42–10, which is concave as viewed from the direction of the incident light, is made convex, the rays will *diverge* after reflection and will form a *virtual* image, as Fig. 42–12 shows. Thus the indented statement above suggests our second sign convention:

2. The radius of curvature r is positive if the center of curvature of the mirror lies on the R-side, as in Fig. 42–10; r is negative if the center of curvature lies on the V-side, as in Fig. 42–12.

The student should not commit these sign conventions to memory but should deduce them in each case from the basic statement on p. 1044, using Fig. 42–10 as a mnemonic aid.

For all cases in this book the object distance o is to be taken as positive. In systems of two or more mirrors (or combinations of mirrors and refracting surfaces) it is possible to arrange that *converging* light falls on the mirror. In such cases the object is called *virtual* and the object distance o is negative; we limit our discussion here to real objects.

▶ **Example 3.** A convex mirror has a radius of curvature of 20 cm. If a point source is placed 14 cm away from the mirror, as in Fig. 42–12, where is the image?

A rough graphical construction, applying the law of reflection at a in the figure, shows that the image will be on the V-side of the mirror and thus will be virtual. We may verify this quantitatively and analytically from Eq. 42–4, noting that r is negative here because the center of curvature of the mirror is on its V-side. We have

$$\frac{1}{o} + \frac{1}{i} = \frac{2}{r}$$

or

$$\frac{1}{+14\text{ cm}} + \frac{1}{i} = \frac{2}{-20\text{ cm}},$$

which yields $i = -5.8$ cm, in agreement with the graphical prediction. The negative sign for i reminds us that the image is on the V-side of the mirror and thus is virtual. ◀

When *parallel* light falls on a mirror (Fig. 42–13), the image point (real or virtual) is called the *focal point* F of the mirror. The focal length f is the distance between F and the vertex. If we put $o \to \infty$ in Eq. 42–4, thus insuring parallel incident light, we have

$$i = \tfrac{1}{2}r = f.$$

---

* This nomenclature may seem unnecessarily cumbersome at this point. We adopt it because of the later extension of these ideas to refracting surfaces and lenses.

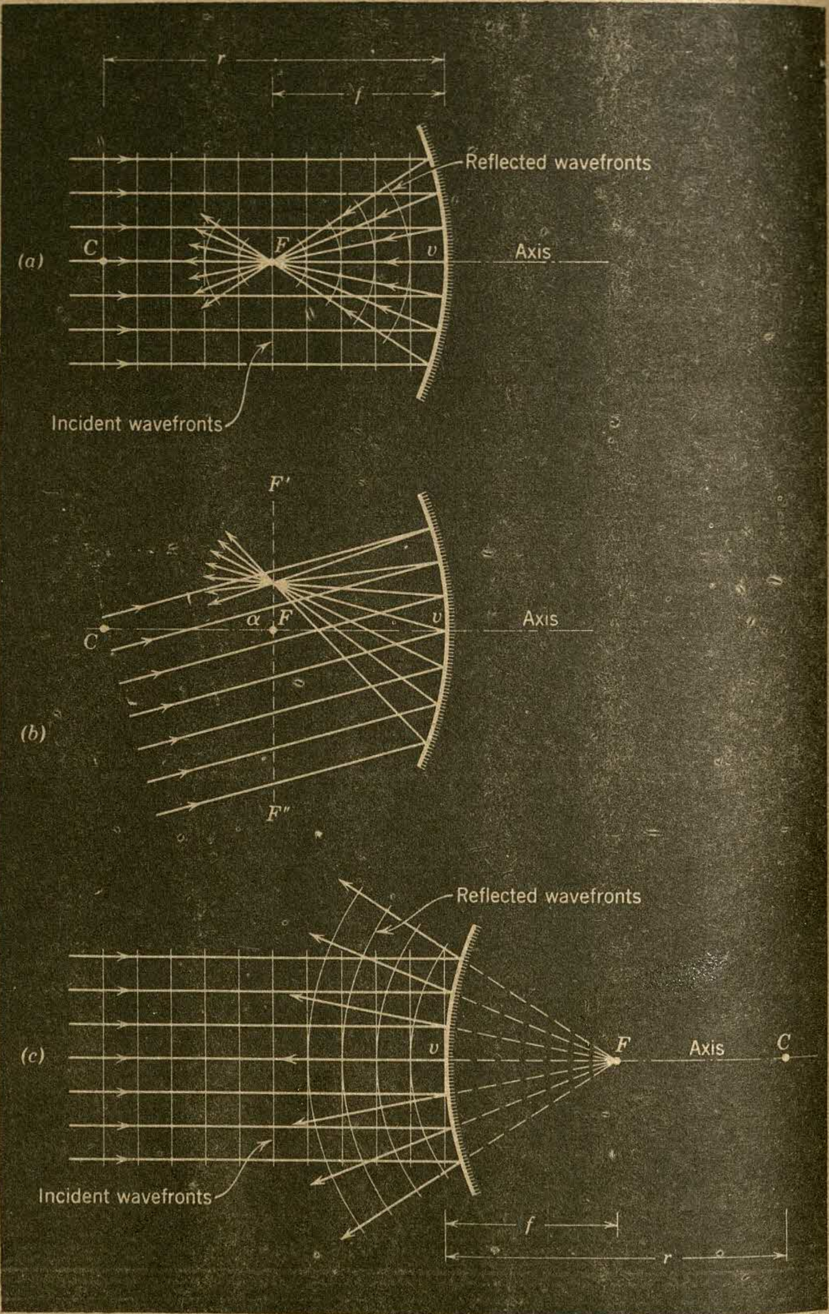


Fig. 42–13 (*a*) The focal point for a concave spherical mirror, showing both the rays and the wavefronts. F and C lie on the R-side, the focal point is real, and the focal length f of the mirror is positive (as is r). (*b*) The same, except that the incident light makes an angle α with the mirror axis; the rays are focused at a point in the *focal plane* $F'F''$. (*c*) Same as (*a*) except that the mirror is convex; F and C lie on the V-side of the mirror. The focal point is virtual and the focal length f is negative (as is r).

Equation 42-4 can then be rewritten

$$\frac{1}{o}+\frac{1}{i}=\frac{1}{f}, \tag{42-5}$$

where f, like r, is taken as positive for mirrors whose centers of curvature are on the R-side (that is, for *concave*, or *converging* mirrors; see Fig. 42-13a) and negative for those whose centers of curvature are on the V-side (that is, for *convex*, or *diverging* mirrors; see Fig. 42-13c). Figure 42-13b shows an incident plane wave that makes a small angle α with the mirror axis. The rays are focused at a point in the *focal plane* of the mirror. This is a plane at right angles to the mirror axis at the focal point.

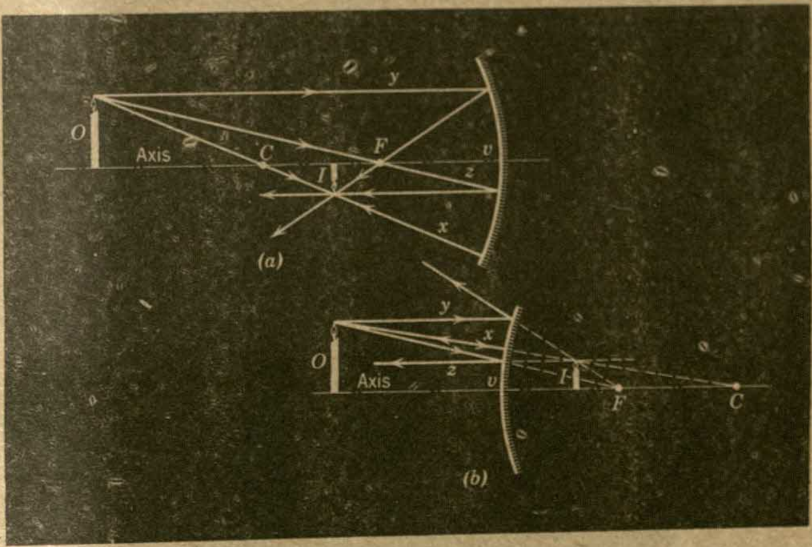


Fig. 42–14 The image of an extended object in (a) a concave mirror and (b) a convex mirror is located graphically. Any two of the three special rays shown are sufficient.

We now consider objects that are not points. Figure 42–14 shows a candle in front of (a) a concave mirror and (b) a convex mirror. We choose to draw the mirror axis through the foot of the candle and, of course, through the center of curvature. The image of any off-axis point, such as the tip of the candle, can be found graphically, using the following facts:

1. A ray that strikes the mirror after passing (either directly or upon being extended) through the center of curvature C returns along itself (ray x in Fig. 42–13). Such rays strike the mirror at right angles.
2. A ray that strikes the mirror parallel to its axis passes (or will pass when extended) through the focal point (ray y).

3. A ray that strikes the mirror after passing (either directly or upon being extended) through the focal point emerges parallel to the axis (ray z).

Figure 42–15 shows a ray (dve) that originates on the tip of the object candle of Fig. 42–14a, is reflected from the mirror at point v, and passes through the tip of the image candle. The law of reflection demands that this ray make equal angles θ with the mirror axis as shown. For the two similar right triangles in the figure we can write

$$\frac{ce}{bd} = \frac{vc}{vb}.$$

The quantity on the left (apart from a question of sign) is the *lateral magnification* m of the mirror. Since we want to represent an *inverted* image by a *negative* magnification, we arbitrarily define m for this case as $-(ce/bd)$. Since $vc = i$ and $vb = o$, we have at once

$$m = -\frac{i}{o}. \qquad (42\text{–}6)$$

This equation gives the magnification for spherical and plane mirrors under all circumstances. For a plane mirror, $o = -i$ and the predicted magnification is $+1$ which, in agreement with experience, indicates an *erect* image the same size as the object.

Images in spherical mirrors suffer from several "defects" that arise because the assumption of paraxial rays is never completely justified. In general, a point source will not produce a point image; see Problem 5. Apart from this, distortion arises because the magnification varies somewhat with distance from the mirror axis, Eq. 42–6 being strictly correct only for paraxial rays. Superimposed on these defects are *diffraction effects* which come about because the basic assumption of geometrical optics, that light travels in straight lines, must always be considered an approximation.

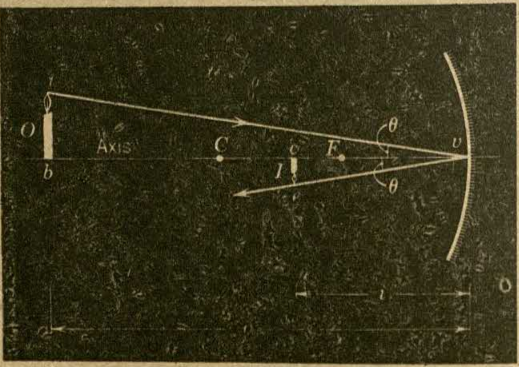


Fig. 42–15 A particular ray for the arrangement of Fig. 42–14, used to show that the *lateral magnification* m is given by $-i/o$.

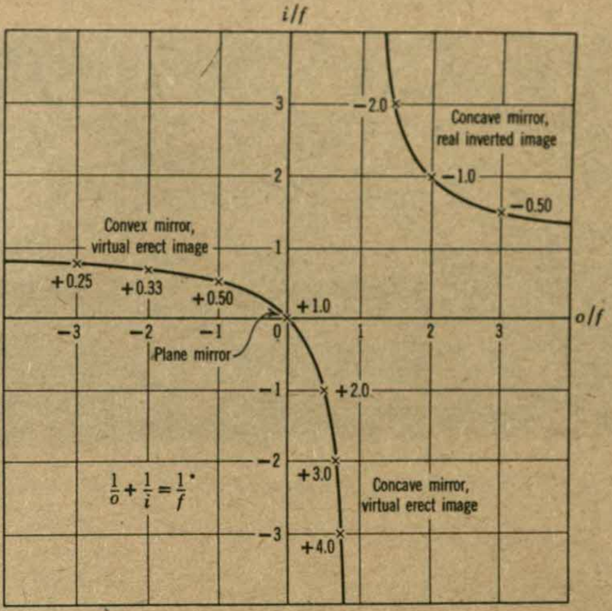


Fig. 42–16 A graphical representation of the mirror formula (Eq. 42–5). The numbers on the curves give the lateral magnifications (Eq. 42–6).

Figure 42–16 summarizes the predictions of the mirror equation (Eq. 42–5) and of the lateral magnification equation (Eq. 42–6). For convex mirrors the image is always virtual. For concave mirrors the image is virtual if the object is inside the focal point ($o/f < 1$) and real otherwise ($o/f > 1$). As $o \to \infty$ for both types of mirror, the image approaches the focal point ($i/f \to 1$). A plane mirror ($m = +1$) is represented in the figure by a point at the origin. It is appropriate that this point form the central point of the lower-left branch of Fig. 42–16 because both convex and concave mirrors approach plane mirrors as r is increased in magnitude, that is, as $f \to \infty$ (or as $o/f \to 0$ in Fig. 42–16, assuming a fixed object distance).

42–4 Spherical Refracting Surface

Figure 42–17 shows a point source O near a convex spherical *refracting* surface of radius of curvature r. The surface separates two media whose indices of refraction differ; *that of the medium in which the incident light falls on the surface being* n_1 and that on the other side of the surface being n_2.

From O we draw a line through the center of curvature C of the refracting surface, thus establishing a convenient axis which intercepts the surface at vertex v. From O we draw a ray that makes a small but arbitrary angle α with the axis and strikes the refracting surface at a, being refracted according to

$$n_1 \sin \theta_1 = n_2 \sin \theta_2.$$

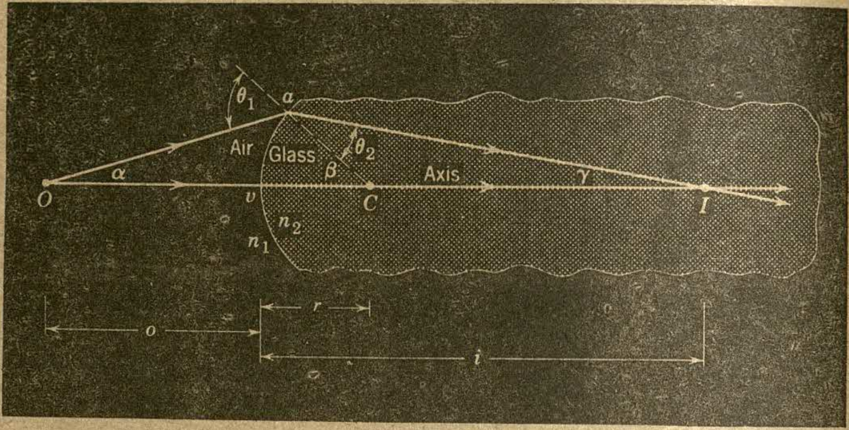


Fig. 42–17 Two rays from O converge after refraction at a spherical surface, forming a real image at I.

The refracted ray intersects the axis at I. A ray from O that travels along the axis will not be bent on entering the surface and will also pass through I. Thus, for these two rays at least, I is the image of O.

As in the derivation of the mirror equation, we use the theorem that the exterior angle of a triangle is equal to the sum of the two opposite interior angles. Applying this to triangles $COaC$ and $ICaI$ yields

$$\theta_1 = \alpha + \beta \tag{42–7}$$

and

$$\beta = \theta_2 + \gamma. \tag{42–8}$$

As α is made small, angles β, γ, θ_1, and θ_2 in Fig. 42–17 also become small. *We at once assume that α, hence all these angles, are arbitrarily small,* this paraxial ray assumption was also made for spherical mirrors. Replacing the sines of the angles by the angles themselves—since the angles are required to be small—permits us to write the law of refraction as

$$n_1\theta_1 \cong n_2\theta_2. \tag{42–9}$$

Combining Eqs. 42–8 and 42–9 leads to

$$\beta = \frac{n_1}{n_2}\theta_1 + \gamma.$$

Eliminating θ_1 between this equation and Eq. 42–7 leads, after rearrangement, to

$$n_1\alpha + n_2\gamma = (n_2 - n_1)\beta. \tag{42–10}$$

In radian measure the angles α, β, and γ in Fig. 42-17 are

$$\alpha \cong \frac{av}{o}$$

$$\beta = \frac{av}{r} \qquad (42\text{-}11)$$

$$\gamma \cong \frac{av}{i}.$$

Only the second of these equations is exact. The other two are approximate because I and O are *not* the centers of circles of which av is an arc. However, for paraxial rays (α small enough) the inaccuracies in Eq. 42-11 can be made as small as desired.

Substituting Eqs. 42-11 into Eq. 42-10 leads readily to

$$\frac{n_1}{o} + \frac{n_2}{i} = \frac{n_2 - n_1}{r}. \qquad (42\text{-}12)$$

This equation holds whenever light is refracted from point objects at spherical surfaces, assuming only that the rays are paraxial. As with the mirror formula, care must be taken to use Eq. 42-12 with consistent signs for o, i, and r. Once again we establish our sign conventions by physical reasoning from a particular case, that of Fig. 42-17:

> In Fig. 42-17, in which light *diverges* from a *real* object, falls on a *convex* refracting surface, and *converges* after refraction to form a *real* image, the quantities o, i, and r in Eq. 42-12 have positive numerical values.

Figure 42-17 was used to derive Eq. 42-12, and the student should associate them in his mind as an aid in getting the signs correct. This basic statement is quite similar to that which we made for mirrors on p. 1044.

We fix our attention on the side of the refracting surface from which the incident light falls on the surface. In contrast to mirrors, the light energy *passes through* a refracting surface to the other side, and if a real image is formed it must appear on the far side, which we call the R-side. The side from which the incident light comes is called the V-side because virtual images must appear here. Figure 42-18 suggests this important distinction between reflection and refraction.

In the indented statement above we associated real images with positive image distances. Thus we are led to the sign convention:

1. The image distance i is positive if the image (real) is on the R-side of the refracting surface, as in Fig. 42-17; i is negative if the image (virtual) lies on the V-side, as in Fig. 42-19.

The refracting surface in Fig. 42-17 is convex. If it is made concave (still assuming that $n_2 > n_1$), the rays will *diverge* after refraction and form

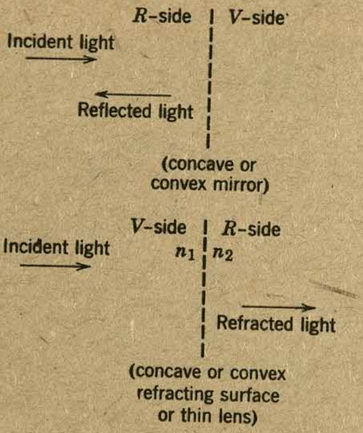


Fig. 42-18 Real images are formed on the same side as the incident light for mirrors but on the opposite side for refracting surfaces and thin lenses. This is so because the incident light is reflected back by mirrors but is transmitted through by refracting surfaces.

a *virtual* image, as Fig. 42-19 shows. Thus we are led to our second sign convention:

2. The radius of curvature r is positive if the center of curvature of the refracting surface lies on the R-side, as in Fig. 42-17; r is negative if the center of curvature is on the V-side, as in Fig. 42-19.

The sign conventions for refracting surfaces are the same as for mirrors (p. 1045), the fundamental difference between the two situations being absorbed in the definitions of R-side and V-side in Fig. 42-18. This difference is easily remembered on physical grounds.

For all cases in this book the object distance o is to be taken as positive. In systems of two or more refracting surfaces (or combinations of refracting surfaces and mirrors) it is possible to arrange that *converging* light falls on the refracting surface. In such cases the object is called *virtual* and the object distance o is negative; we limit our discussions here to real objects.

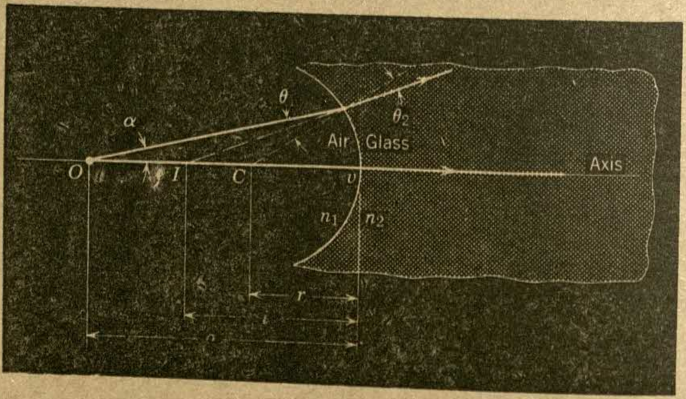


Fig. 42-19 Two rays from O diverge after refraction at a spherical surface, forming a virtual image at I.

▶ **Example 4.** Locate the image for the geometry shown in Fig. 42-17, assuming the radius of curvature to be 10 cm, n_2 to be 2.0, and n_1 to be 1.0. Let the object be 20 cm to the left of v.

From Eq. 42-12,

$$\frac{n_1}{o} + \frac{n_2}{i} = \frac{n_2 - n_1}{r},$$

we have

$$\frac{1.0}{+20\text{ cm}} + \frac{2.0}{i} = \frac{2.0 - 1.0}{+10\text{ cm}}.$$

Note that r is positive because the center of curvature of the surface lies on the R-side. This relation yields $i = +40$ cm in agreement with the graphical construction. The light energy actually passes through I so that the image is real, as indicated by the positive sign for i.

Example 5. An object is immersed in a medium with $n_1 = 2.0$, being 15 cm from the spherical surface whose radius of curvature is -10 cm, as in Fig. 42-20; r is negative because C lies on the V-side. Locate the image.

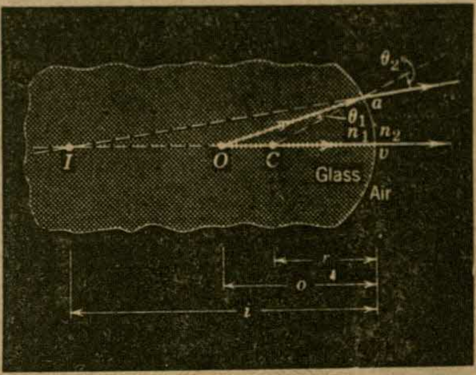


Fig. 42-20 Two rays from O appear to originate from I (virtual image) after refraction at a spherical surface.

Figure 42-20 shows a ray traced through the surface by applying the law of refraction at point a. A second ray from O along the axis emerges undeflected at v. The image I is found by extending these two rays backward; it is virtual.

From Eq. 42-12,

$$\frac{n_1}{o} + \frac{n_2}{i} = \frac{n_2 - n_1}{r},$$

we have

$$\frac{2.0}{+15\text{ cm}} + \frac{1.0}{i} = \frac{1.0 - 2.0}{-10\text{ cm}},$$

which yields $i = -30$ cm, in agreement with Fig. 42-20 and with the sign conventions. Note that n_1 always refers to the medium on the side of the surface from which the light comes.

Example 6. What is the relationship between i and o if the refracting surface is plane?

A plane surface has an infinite radius of curvature. Putting $r \to \infty$ in Eq. 42-12 leads to

$$i = -o\frac{n_2}{n_1}.$$

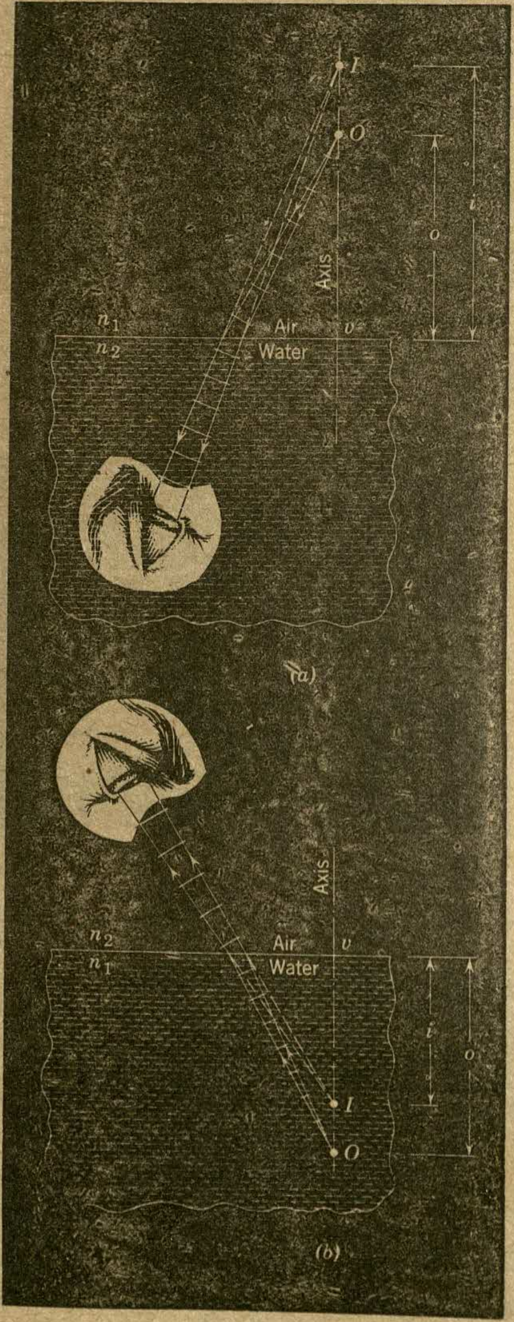


Fig. 42–21 Refraction at a plane surface at near-normal incidence, showing a pencil of rays and the corresponding wavefronts entering the pupil. (*a*) Source in air and (*b*) source in water.

Figure 42-21 illustrates the situation graphically (*a*) for an object in air as seen from below water and (*b*) for an object in water with air above. This shows that a diver, looking upward at, say, an overhanging tree branch, will think it higher than it is by the factor 1.33/1.00. Similarly, an observer in air will think that objects on the bottom of a water tank are closer to the surface than they actually are, in the ratio 1.00/1.33. These considerations, being based on Eq. 42-12, hold only for paraxial rays, which means that the incident rays can make only a small angle with the normal; this angle has been exaggerated in the figure for clarity. Note again that n_1 is always identified with the medium that lies on the side of the surface containing the incident light.

42-5 Thin Lenses

In most refraction situations there is more than one refracting surface. This is true even for a spectacle lens, the light passing from air into glass and then from glass into air. In microscopes, telescopes, cameras, etc., there are often many more than two surfaces.

Figure 42-22*a* shows a thick glass "lens" of length l whose surfaces are ground to radii r' and r''. A point object O' is placed near the left surface as shown. A ray leaving O' along the axis is not deflected on entering or leaving the lens because it falls on each surface along a normal.

A second ray leaving O', at an arbitrary angle α with the axis, strikes the surface at point a', is refracted, and strikes the second surface at point a''. The ray is again refracted and crosses the axis at I'', which, being the intersection of two rays from O'', is the image of point O', formed after refraction at two surfaces.

Figure 42-22*b* shows the first surface, which forms a virtual image of O' at I'. To locate I', we use Eq. 42-12,

$$\frac{n_1}{o} + \frac{n_2}{i} = \frac{n_2 - n_1}{r}.$$

Putting $n_1 = 1.0$ and $n_2 = n$ and bearing in mind that the image distance is negative (that is, $i = -i'$ in Fig. 42-22*b*), we obtain

$$\frac{1}{o'} - \frac{n}{i'} = \frac{n-1}{r'}. \qquad (42\text{-}13)$$

In this equation i' will be a positive number because we have arbitrarily introduced the minus sign appropriate to a virtual image.

Figure 42-22*c* shows the second surface. Unless an observer at point a'' were aware of the existence of the first surface, he would think that the light striking that point originated at point I' in Fig. 42-22*b* and that the region to the left of the surface was filled with glass. Thus the (virtual) image I' formed by the first surface serves as a real object O'' for the second surface. The distance of this object from the second surface is

$$o'' = i' + l. \qquad (42\text{-}14)$$

In applying Eq. 42-12 to the second surface, we insert $n_1 = n$ and $n_2 = 1.0$ because the object behaves as if it were imbedded in glass. If we use

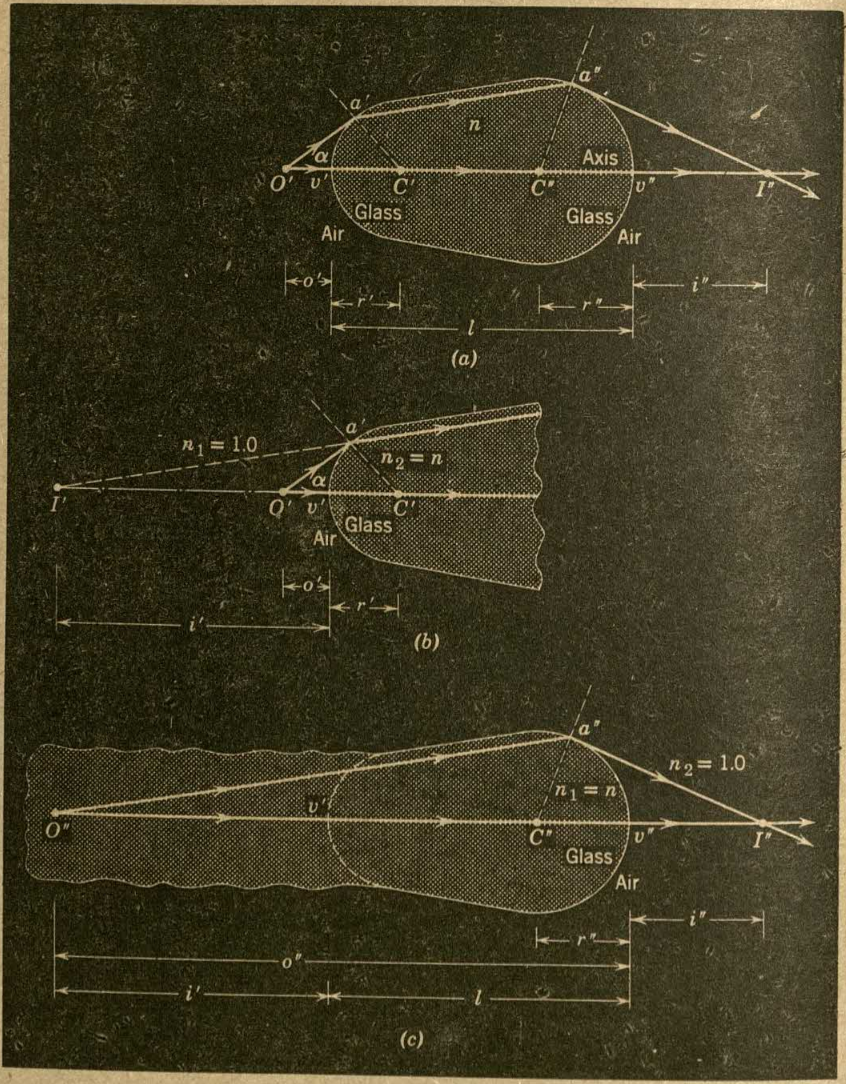


Fig. 42–22 (*a*) Two rays from O' intersect at I'' (real image) after refraction at two spherical surfaces. (*b*) The first surface and (*c*) the second surface shown separately. The quantities α and n have been exaggerated for clarity.

Eq. 42–14, Eq. 42–12 becomes

$$\frac{n}{i' + l} + \frac{1}{i''} = \frac{1 - n}{r''}. \tag{42–15}$$

Let us now assume that the thickness l of the "lens" in Fig. 42–22 is so small that it can be neglected in comparison with other linear quantities in this figure (such as o', i', o'', i'', r', and r''). In all that follows we make this *thin-lens approximation*. Putting $l = 0$ in Eq. 42–15 leads to

$$\frac{n}{i'} + \frac{1}{i''} = -\frac{n-1}{r''}. \tag{42–16}$$

Adding Eqs. 42–13 and 42–16 leads to

$$\frac{1}{o'} + \frac{1}{i''} = (n-1)\left(\frac{1}{r'} - \frac{1}{r''}\right).$$

Finally, calling the original object distance simply o and the final image distance simply i leads to

$$\frac{1}{o} + \frac{1}{i} = (n-1)\left(\frac{1}{r'} - \frac{1}{r''}\right). \tag{42–17}$$

This equation holds only for paraxial rays and only if the lens is so thin that it essentially makes no difference from which surface of the lens the quantities o and i are measured. In Eq. 42–17 r' refers to the first surface struck by the light as it traverses the lens and r'' to the second surface.

The sign conventions for Eq. 42–17 are the same as those for mirrors and for single refracting surfaces. Because the lens is assumed to be thin, we refer to the R-side and the V-side of the lens itself (see Fig. 42–18) rather than those of its separate surfaces. The sign conventions then are the following:

1. The image distance i is positive if the image (real) lies on the R-side of the lens, as in Fig. 42–23*a*; i is negative if the image (virtual) lies on the V-side of the lens, as in Fig. 42–23*b*.
2. The radii of curvature r' and r'' are positive if their respective centers

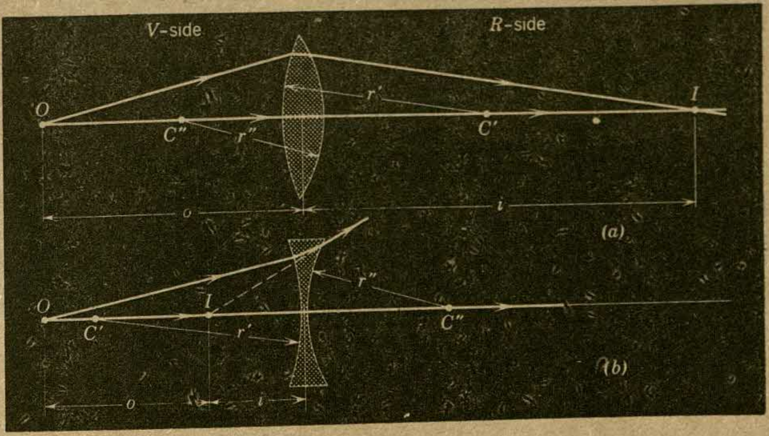


Fig. 42–23 Illustrations used to establish sign conventions for thin lenses.

of curvature lie on the R-side of the lens. They are negative if their centers lie on the V-side. In Fig. 42–23*a*, r' is positive and r'' is negative; in Fig. 42–23*b*, r' is negative and r'' is positive.

Figure 42–24*a* and *c* shows parallel light from a distant object falling on a thin lens. The image location is called the *second focal point* F_2 of the lens. The distance from F_2 to the lens is called the *focal length* f. The *first focal point* for a thin lens (F_1 in figure) is the object position for which the image

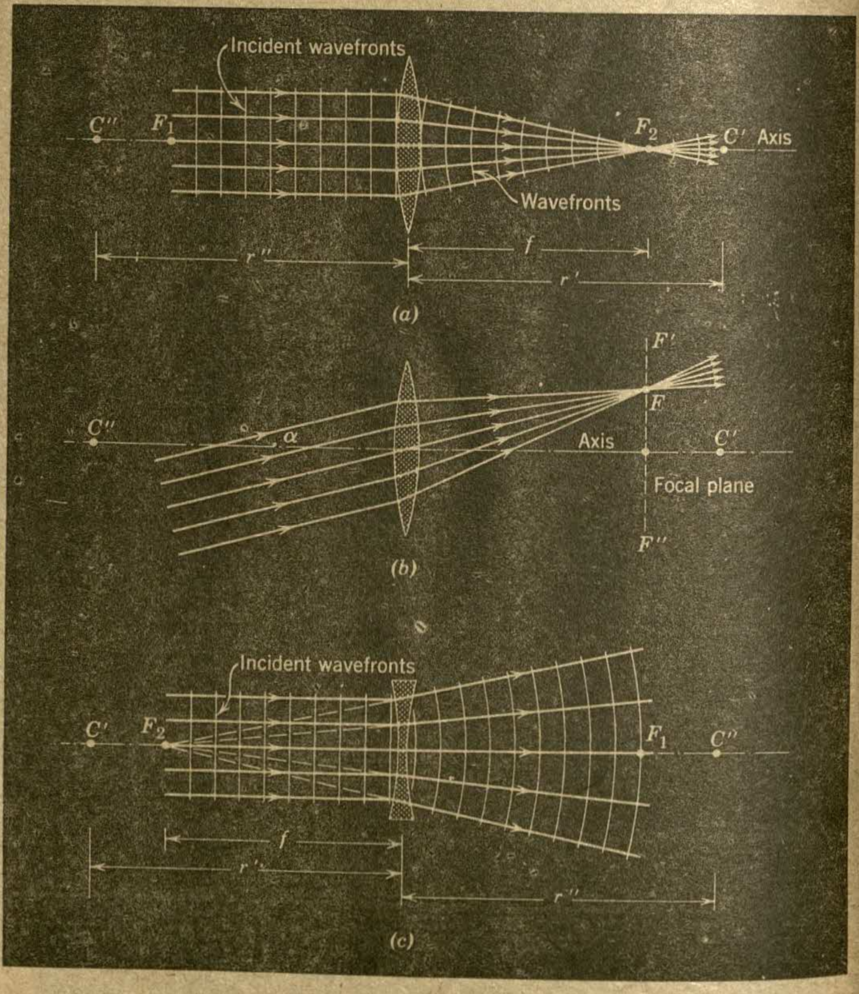


Fig. 42–24 (*a*) Parallel light passes through the second focal point F_2 of a converging lens. (*b*) The incident light makes an angle α with the lens axis, the rays being focused in the focal plane $F'F''$. (*c*) Parallel light, passing through a diverging lens, seems to originate at the second focal point F_2. C' and C'' are centers of curvature for the lens surfaces; F_1 is the first focal point.

is at infinity. For thin lenses the first and second focal points are on opposite sides of the lens and are equidistant from it.

The focal length can be computed from Eq. 42-17 by inserting $o \to \infty$ and $i = f$. This yields

$$\frac{1}{f} = (n - 1)\left(\frac{1}{r'} - \frac{1}{r''}\right). \tag{42-18}$$

This relation is called the *lens maker's equation* because it allows us to compute the focal length of a lens in terms of the radii of curvature and the index of refraction of the material. Combining Eqs. 42-17 and 42-18 allows us to write the thin-lens equation as

$$\frac{1}{o} + \frac{1}{i} = \frac{1}{f}. \tag{42-19}$$

Figure 42-24*b* shows parallel incident rays that make a small angle α with the lens axis; they are brought to a focus in the *focal plane* $F'F''$, as shown. This is a plane normal to the lens axis at the focal point.

In Fig. 42-24*a* we note that all rays in the figure contain the same number of wavelengths; in other words, they have the same *optical path lengths;* see Section 41-6. The optical path lengths are the same because the wavefronts are surfaces over which the wave disturbance has the same constant value and because all the rays shown pass through the same number of wavefronts.

▶ **Example 7.** The lenses of Fig. 42-24 have radii of curvature of magnitude 40 cm and are made of glass with $n = 1.65$. Compute their focal lengths.

Since C' lies on the R-side of the lens in Fig. 42-24*a*, r' is positive (= +40 cm). Since C'' lies on the V-side, r'' is negative (= −40 cm). Substituting in Eq. 42-18 yields

$$\frac{1}{f} = (n - 1)\left(\frac{1}{r'} - \frac{1}{r''}\right) = (1.65 - 1)\left(\frac{1}{+40\text{ cm}} - \frac{1}{-40\text{ cm}}\right),$$

or

$$f = +31\text{ cm}.$$

A positive focal length indicates that in agreement with Fig. 42-24*a* the focal point F_2 is on the R-side of the lens and parallel incident light converges after refraction to form a real image.

In Fig. 42-24*c* C' lies on the V-side of the lens so that r' is negative (= −40 cm). Since r'' is positive (= +40 cm), Eq. 42-17 yields

$$f = -31\text{ cm}.$$

A negative focal length indicates that in agreement with Fig. 42-24*c* the focal point F_2 is on the V-side of the lens and incident light diverges after refraction to form a virtual image ◀

The location of the image of an extended object such as a candle (Fig. 42-25) can be found graphically by using the following three facts:

1. A ray parallel to the axis and falling on the lens passes, either directly or when extended, through the second focal point (ray x in Fig. 42-25).
2. A ray falling on a lens after passing, either directly or when extended,

through the first focal point will emerge from the lens parallel to the axis (ray y).

3. A ray falling on the lens at its center will pass through undeflected. There is no deflection because the lens, near its center, behaves like a thin piece of glass with parallel sides. The direction of the light rays is not changed and the sideways displacement can be neglected because the lens thickness has been assumed to be negligible (ray z; see also Problem 13, Chapter 41).

Figure 42–26, which represents part of Fig. 42–25a, shows a ray passing from the tip of the object through the center of curvature to the tip of the image. For the similar triangles abc and dec we may write

$$\frac{de}{ab} = \frac{dc}{ac}.$$

The right side of this equation is i/o and the left side is $-m$, where m is the

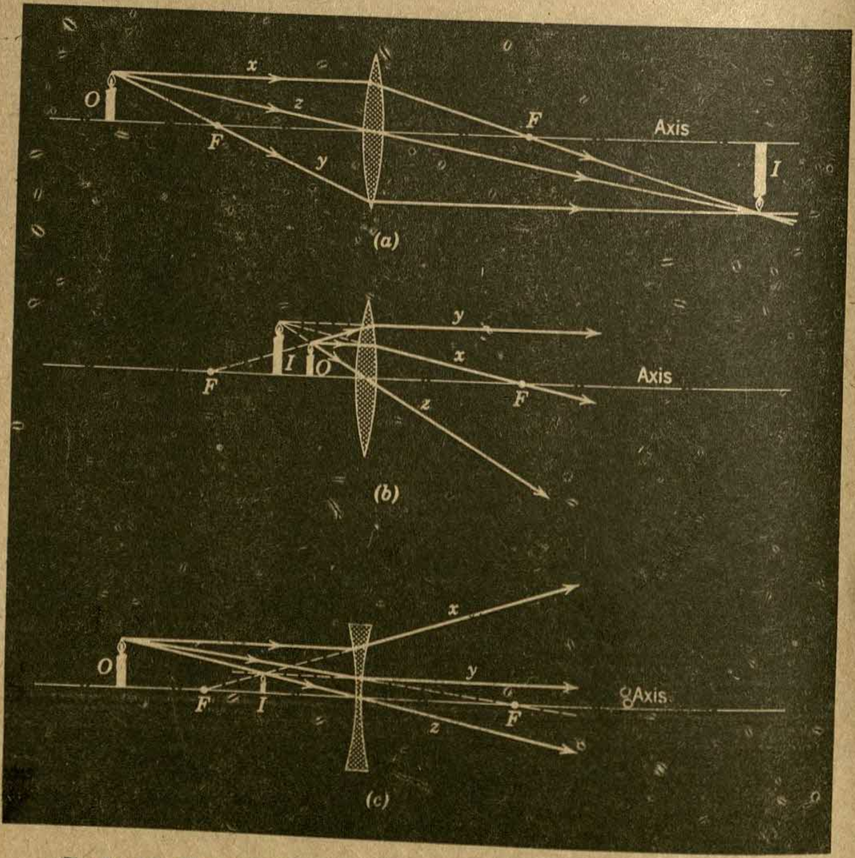


Fig. 42–25 Showing the graphical location of images for three thin lenses.

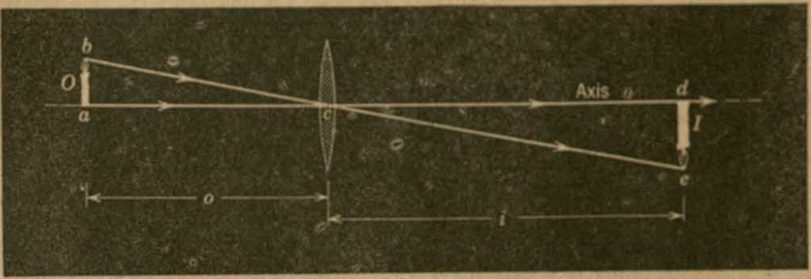


Fig. 42-26 Two rays for the situation of Fig. 42-25*a*.

lateral magnification. The minus sign is required because we wish m to be negative for an inverted image. This yields

$$m = -\frac{i}{o}, \qquad (42\text{-}20)$$

which holds for all types of thin lenses and for all object distances.

▶ **Example 8.** A converging thin lens has a focal length of +24 cm. An object is placed 9.0 cm from the lens as in Fig. 42-25*b*; describe the image.

From Eq. 42-19,

$$\frac{1}{o} + \frac{1}{i} = \frac{1}{f},$$

we have

$$\frac{1}{+9.0\ \text{cm}} + \frac{1}{i} = \frac{1}{+24\ \text{cm}},$$

which yields $i = -14.4$ cm, in agreement with the figure. The minus sign means that the image is on the V-side of the lens and is thus virtual.

The lateral magnification is given by

$$m = -\frac{i}{o} = -\frac{-14.4\ \text{cm}}{+9.0\ \text{cm}} = +1.6,$$

again in agreement with the figure. The plus signifies an erect image. ◀

Images formed by lenses suffer from defects similar to those discussed for mirrors on p. 1048. There are effects connected with the failure of a point object to form a point image, with the variation of magnification with distance from the lens axis, and with diffraction. For lenses, but not for mirrors, there are also *chromatic aberrations* associated with the fact that the refracting properties of the lens vary with wavelength because the index of refraction of the lens material does. If a point object on the lens axis emits white light, the image, neglecting other lens defects, will be a series of colored points spread out along the axis. We have all seen the colored images produced by inexpensive lenses. A great deal of ingenious op[illegible]gineering goes into the design of lenses (more commonly lens syste[illegible] the various lens defects are minimized. The lens surfaces a[illegible] spherical.

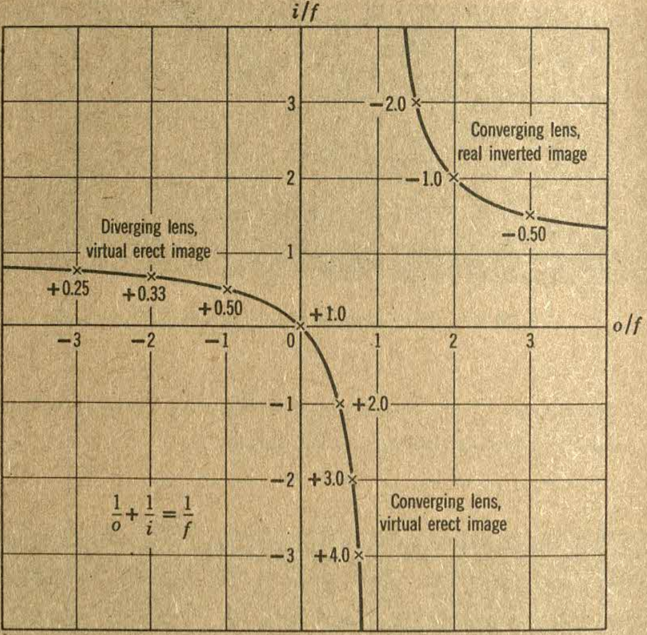


Fig. 42–27 A graphical representation of the thin-lens formula (Eq. 42–19). The numbers on the curves give the lateral magnifications (Eq. 42–20). Compare Fig. 42–16.

Figure 42–27 is a graphical representation of the object-image relationships for thin lenses; see Eqs. 42–19 and 42–20. The fact that it is identical in form with Fig. 42–16, which holds for mirrors, is not surprising because the basic equations for mirrors and thin lenses (Eqs. 42–19 and 42–5; Eqs. 42–20 and 42–6) are identical. A converging lens of a given focal length is quantitatively similar to a concave mirror of the same focal length, the difference being that the lens forms real images on the opposite side from the (real) object and the mirror forms them on the same side.

QUESTIONS

1. If a mirror reverses right and left, why doesn't it reverse up and down?
2. Is it possible to photograph a virtual image?
3. What approximations were made in deriving the mirror equation (Eq. 42–4):

$$\frac{1}{o} + \frac{1}{i} = \frac{2}{r}?$$

4. Can you think of a simple test or observation to prove that the law of reflection is the same for all wavelengths, under conditions in which geometrical optics prevails?
5. Under what conditions will a spherical mirror, which may be concave or convex, form (*a*) a real image, (*b*) an inverted image, and (*c*) an image smaller than the object?
6. An unsymmetrical thin lens forms an image of a point object on its axis. Is the image location changed if the lens is reversed?
7. Why has a lens two focal points and a mirror only one?

8. Under what conditions will a thin lens, which may be converging or diverging, form (*a*) a real image, (*b*) an inverted image, and (*c*) an image smaller than the object?

9. Does the apparent depth of an object below water depend on the angle of view of the observer in air? Explain and illustrate with ray diagrams.

10. A skin diver wants to use an air-filled plastic bag as a converging lens for underwater use. Sketch a suitable cross section for the bag.

11. What approximations were made in deriving the thin lens equation (Eq. 42–19):

$$\frac{1}{o}+\frac{1}{i}=\frac{1}{f}?$$

12. Under what conditions will a thin lens have a lateral magnification (*a*) of -1 and (*b*) of $+1$?

13. How does the focal length of a glass lens for blue light compare with that for red light, assuming the lens is (*a*) diverging and (*b*) converging?

14. Does the focal length of a lens depend on the medium in which the lens is immersed? Is it possible for a given lens to act as a converging lens in one medium and a diverging lens in another medium?

15. Are the following statements true for a glass lens in air? (*a*) A lens that is thicker at the center than at the edges is a converging lens for parallel light. (*b*) A lens that is thicker at the edges than at the center is a diverging lens for parallel light. Explain and illustrate, using wavefronts.

16. Under what conditions would the lateral magnification, ($m = -i/o$) for lenses and mirrors become infinite? Is there any practical significance to such a condition?

17. Light rays are reversible. Discuss the situation in terms of objects and images if all rays in Figs. 42–10, 42–14, 42–17, 42–19, 42–23, and 42–25 are reversed in direction.

18. What significance can be given to the origin of coordinates in the graphical representation of the thin-lens formula (Fig. 42–27)?

19. In connection with Fig. 42–24*a*, we pointed out that all rays originating on the same wavefront in the incident wave have the same optical path length to the image point. Discuss this in connection with Fermat's principle (Section 41–6).

PROBLEMS

1. Solve Example 2 if the angle between the mirrors is (*a*) 45°, (*b*) 60°, (*c*) 120°, the object always being placed on the bisector of the mirrors.

2. Two plane mirrors make an angle of 90° with each other. What is the largest number of images of an object placed between them that can be seen by a properly placed eye? The object need *not* lie on the mirror bisector.

3. A small object is 10 cm in front of a plane mirror. If you stand behind the object, 30 cm from the mirror, and look at its image, for what distance must you focus your eyes?

4. A small object O is placed one-third of the way between two parallel plane mirrors as in Fig. 42–28. Trace appropriate bundles of rays for viewing the four images that lie closest to the object.

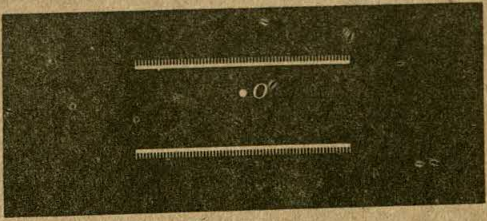


Fig. 42–28

5. Redraw Fig. 42–29 on a large sheet of paper and trace carefully the reflected rays, using the law of reflection. Is a point focus formed? Discuss.

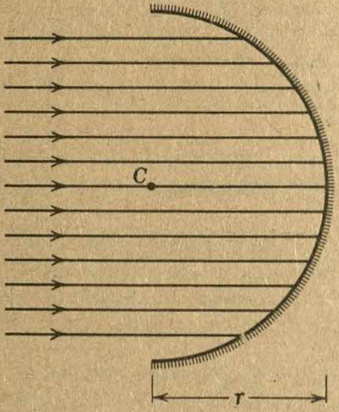


Fig. 42–29

6. Fill in this table, each column of which refers to a spherical mirror. Check your results by graphical analysis. Distances are in centimeters; if a number has no plus or minus sign in front of it, it may have either sign.

	a	*b*	*c*	*d*	*e*	*f*	*g*	*h*
Type	concave						convex	
f	20		+20			20		
r					−40		40	
i					−10		4	
o	+10	+10	+30	+60				+24
m		+1		−0.5		+0.10		0.50
Real image?		no						
Erect image?								no

7. A short linear object of length l lies on the axis of a spherical mirror, a distance o from the mirror. (*a*) Show that its image will have a length l' where

$$l' = l\left(\frac{f}{o - f}\right)^2.$$

(*b*) Show that the *longitudinal magnification* $m'(= l'/l)$ is equal to m^2 where m is the lateral magnification discussed in Section 42–3. (*c*) Is there any condition such that, neglecting all mirror defects, the image of a small cube would also be a cube?

8. As an example of the effect of relaxing the paraxial ray assumption find, analytically,

the apparent depth of a swimming pool having water ($n = 1.33$) of actual depth 8.0 ft (*a*) when viewed at normal incidence and (*b*) when viewed at an angle of 30° from the normal.

9. A layer of water ($n = 1.33$) 2.0 cm thick floats on carbon tetrachloride ($n = 1.46$) 4.0 cm thick. How far below the water surface, viewed at normal incidence, does the bottom of the tank seem to be?

10. Fill out the following table, each column of which refers to a spherical surface separating two media with different indices of refraction. Distances are measured in centimeters.

	a	*b*	*c*	*d*	*e*	*f*	*g*	*h*
n_1	1.0	1.0	1.0	1.0	1.5	1.5	1.5	1.5
n_2	1.5	1.5	1.5		1.0	1.0	1.0	
o	+10	+10		+20	+10		+70	+100
i		−13	+600	−20	−6	−7.5		+600
r	+30		+30	−20		−30	+30	−30
Real image?								

Draw a figure for each situation and construct the appropriate rays graphically. Assume a point object.

11. Fill in this table, each column of which refers to a thin lens, to the extent possible. Check your results by graphical analysis. Distances are in centimeters; if a number (except in row n) has no plus sign or minus sign in front of it, it may have either sign.

	a	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*
Type	converging								
f	10	+10	10	10					
r'					+30	−30	−30		
r''					−30	+30	−60		
i									
o	+20	+5	+5	+5	+10	+10	+10	+10	+10
n					1.5	1.5	1.5		
m			>1	<1				0.5	0.5
Real image?									yes
Erect image?								yes	

Draw a figure for each situation and construct the appropriate rays graphically. Assume a finite object.

12. Define and locate the first and second focal points (see p. 966) for a single spherical refracting surface such as that of Fig. 42–17.

13. A luminous object and a screen are a fixed distance D apart. (*a*) Show that a converging lens of focal length f will form a real image on the screen for two positions that are separated by

$$d = \sqrt{D(D - 4f)}.$$

(*b*) Show that the ratio of the two image sizes for these two positions is

$$\left(\frac{D - d}{D + d}\right)^2$$

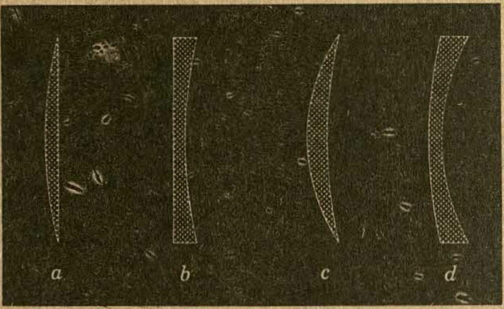


Fig. 42–30

14. Using the lens maker's equation (Eq. 42–18), decide which of the thin lenses in Fig. 42–30 is converging and which diverging for parallel incident light.

15. Two thin lenses of focal length f_1 and f_2 are in contact. Show that they are equivalent to a single thin lens with a focal length given by

$$f = \frac{f_1 f_2}{f_1 + f_2}.$$

16. Show that the distance between an object and its real image formed by a thin converging lens is always greater than four times the focal length of the lens.

17. A double-convex lens is to be made of glass with an index of refraction of 1.50. One surface is to have twice the radius of curvature of the other and the focal length is to be 6.0 cm. What are the radii?

18. The formula

$$\frac{1}{o} + \frac{1}{i} = \frac{1}{f}$$

is called the *Gaussian form* of the thin lens formula. Another form of this formula, the *Newtonian* form, is obtained by considering the distance x from the object to the first focal point and the distance x' from the second focal point to the image. Show that

$$xx' = f^2.$$

19. A parallel incident beam falls on a solid glass sphere at normal incidence. Locate the image in terms of the index of refraction n and the sphere radius r.

20. An erect object is placed a distance in front of a converging lens equal to twice the focal length f_1 of the lens. On the other side of the lens is a converging mirror of focal length f_2 separated from the lens by a distance $2(f_1 + f_2)$. (*a*) Find the location, nature and relative size of the final image. (*b*) Draw the appropriate ray diagram. See Fig. 42–31.

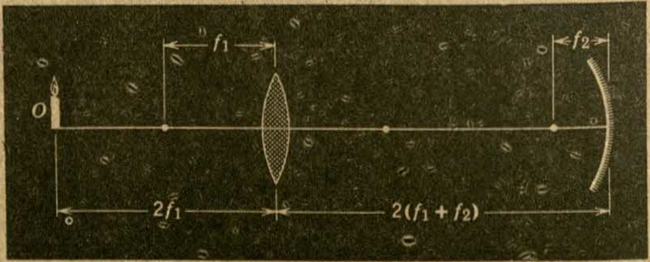


Fig. 42–31

21. An object is placed at a center of curvature of a double-concave lens, both of whose radii of curvature are the same. (*a*) According to the convention used in this book, what are the signs of the two radii of curvature? (*b*) Find the location of the image in terms of the radius of curvature r and the index of refraction, n, of the glass. (*c*) Describe the nature of the image. (*d*) Verify your result with a ray diagram.

22. Two lenses with their focal points are shown in Fig. 42-32. With the given object, find the image formed by lens 1 and, using this as an object for lens 2, find the final image. Discuss the nature of the two images. Verify your conclusions by copying the figure to scale and drawing a ray diagram.

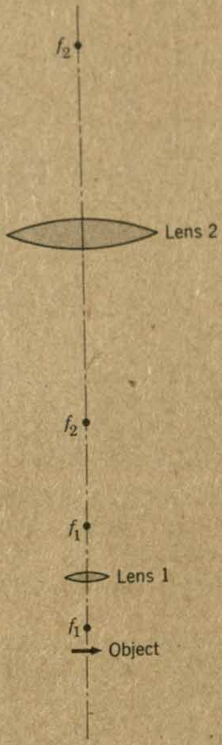


Fig. 42–32

23. An object is 20 cm to the left of a lens with a focal length of +10 cm. A second lens of focal length +12.5 cm is 30 cm to the right of the first lens. (*a*) Using the image formed by the first lens as the object for the second, find the location and relative size of the final image. (*b*) Verify your conclusions by drawing the lens system to scale and constructing a ray diagram. (*c*) Describe the final image.

Interference

CHAPTER 43

43–1 Young's Experiment

In Section 19–7 we saw that if two waves of the same frequency travel in approximately the same direction and have a phase difference that remains constant with time, they may combine so that their energy is not distributed uniformly in space but is a maximum at certain points and a minimum (possibly even zero) at others. The demonstration of such *interference* effects for light by Thomas Young in 1801 first established the wave theory of light on a firm experimental basis. Young was able to deduce the wavelength of light from his experiments, the first measurement of this important quantity.

Young allowed sunlight to fall on a pinhole S_0 punched in a screen A in Fig. 43–1. The emerging light spreads out by diffraction (see Section 42–1) and falls on pinholes S_1 and S_2 punched into screen B. Again diffraction occurs and two overlapping spherical waves expand into the space to the right of screen B.

The condition for geometric optics, namely that $a \gg \lambda$ where a is the diameter of the pinholes, is definitely *not* met in this experiment. The pinholes do not cast geometrical shadows but act as sources of expanding Huygens' wavelets. We are dealing here (and in succeeding chapters) with *wave optics* rather than with geometrical optics.

Figure 43–2, taken from an 1803 paper of Young, shows the region between screens B and C. The blackening represents the minima of the wave disturbance; the white space between represents the maxima. If you hold the page with your eye close to the left edge and look at a grazing angle along the figure, you will see that along lines marked by x's there is cancella-

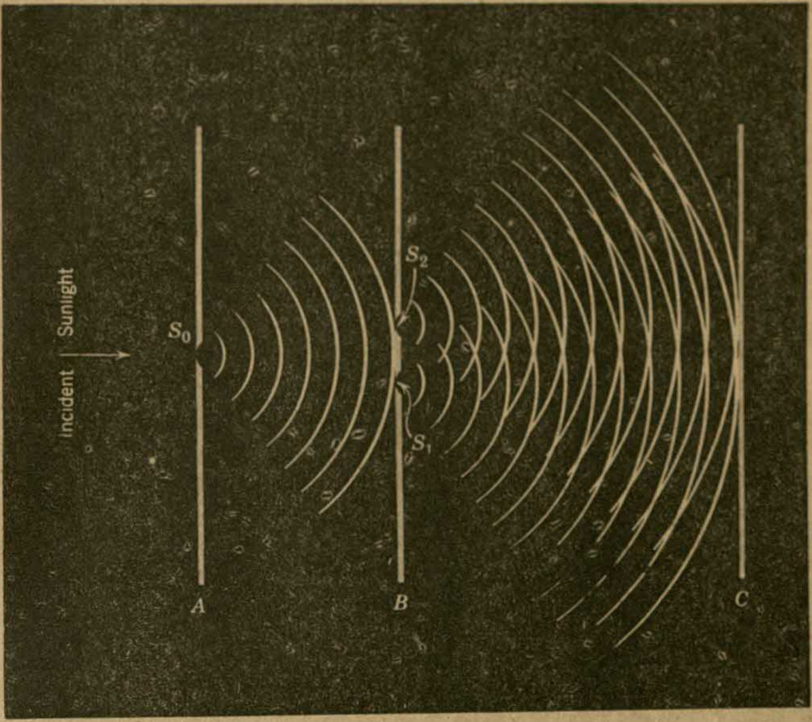


Fig. 43–1 Showing how Thomas Young produced an interference pattern by allowing diffracted waves from pinholes S_1 and S_2 to overlap on screen C.

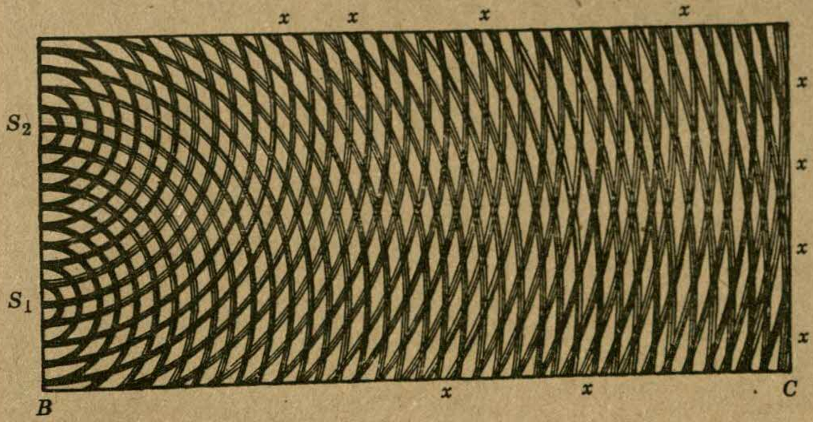


Fig. 43–2 Thomas Young's original drawing showing interference effects in overlapping waves. Place the eye near the left edge and sight at a grazing angle along the figure (From Thomas Young, *Phil. Transactions*, 1803.)

tion of the wave; between them there is reinforcement. If a screen is placed anywhere across the superimposed waves, we expect to find alternate bright and dark spots on it. Figure 43–3 shows a photograph of such *interference fringes;* in keeping with modern technique, long narrow slits rather than pinholes were used in preparing this figure.

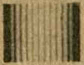


Fig. 43–3 Interference fringes for monochromatic light, made with an arrangement like that of Fig. 43–1, using long narrow slits rather than pinholes.

Interference is not limited to light waves but is a characteristic of all wave phenomena. Figure 43–4, for example, shows the interference pattern of

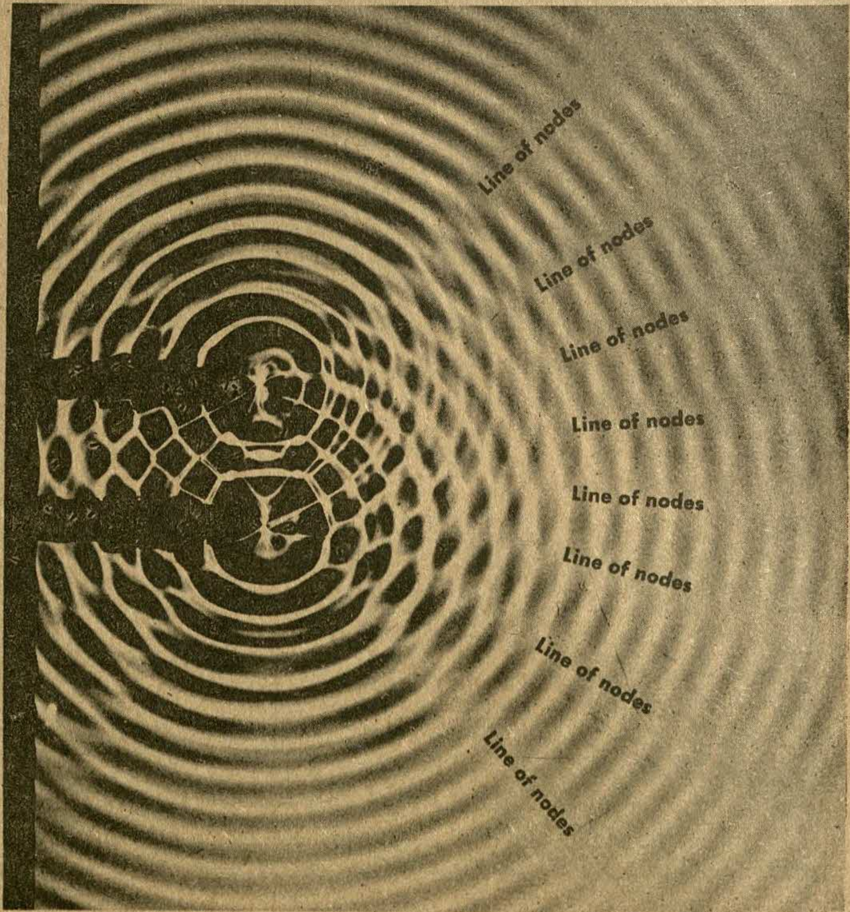


Fig. 43–4 The interference of water waves in a ripple tank. There is destructive interference along the lines marked "Line of nodes" and constructive interference between these lines. (Courtesy Physical Science Study Committee.)

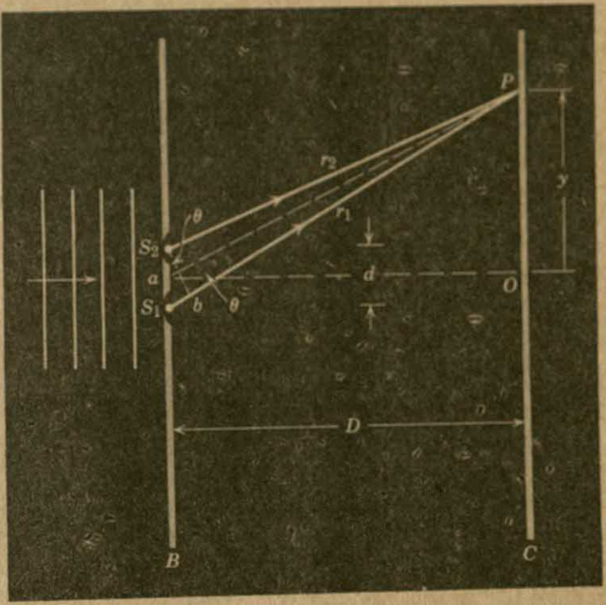


Fig. 43–5 Rays from S_1 and S_2 combine at P. The light falling on screen B has been taken as parallel. Actually, $D \gg d$, the figure being distorted for clarity.

water waves in a shallow *ripple tank*. The waves are generated by two vibrators that tap the water surface in synchronism, producing two expanding spherical waves.

Let us analyze Young's experiment quantitatively, assuming that the incident light consists of a single wavelength only. In Fig. 43–5 P is an arbitrary point on the screen, a distance r_1 and r_2 from the narrow slits S_1 and S_2, respectively. Let us draw a line from S_2 to b in such a way that the lines PS_2 and Pb are equal. If d, the slit spacing, is much smaller than the distance D between the two screens (the ratio d/D in the figure has been exaggerated for clarity), S_2b is then almost perpendicular to both r_1 and r_2. This means that angle S_1S_2b is almost equal to angle PaO, both angles being marked θ in the figure. This is equivalent to saying that the lines r_1 and r_2 may be taken as parallel.

We often put a lens in front of the two slits, as in Fig. 43–6, the screen C being in the focal plane of the lens. Under these conditions light focused at P must have struck the lens parallel to the line Px, drawn from P through the center of the (thin) lens. Under these conditions rays r_1 and r_2 are strictly parallel even though the requirement $D \gg d$ is not met. The lens L may in practice be the lens and cornea of the eye, screen C being the retina.

The two rays arriving at P in Figs. 43–5 or 43–6 from S_1 and S_2 are in phase at the source slits, both being derived from the same wavefront in the incident plane wave. Because the rays have different optical path lengths,

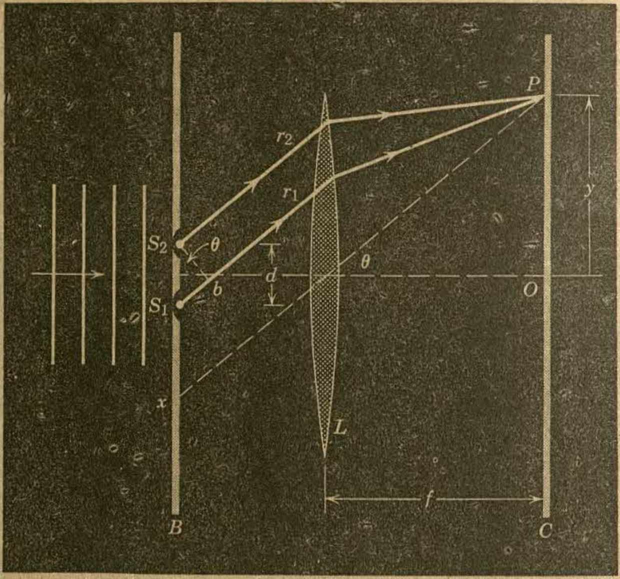


Fig. 43–6 A lens is normally used to produce interference fringes; compare with Fig. 43–5. The figure is again distorted for clarity in that $f \gg d$ in practice.

they arrive at P with a phase difference. The number of wavelengths contained in S_1b, which is the path difference, determines the nature of the interference at P.

To have a *maximum* at P, S_1b $(= d \sin \theta)$ must contain an integral number of wavelengths, or

$$S_1b = m\lambda \qquad m = 0, 1, 2, \ldots,$$

which can be written as

$$d \sin \theta = m\lambda \qquad m = 0, 1, 2, \ldots \text{ (maxima).} \tag{43–1}$$

Note that each maximum above O in Figs. 43–5 and 43–6 has a symmetrically located maximum below O. There is a central maximum described by $m = 0$.

For a *minimum* at P, S_1b $(= d \sin \theta)$ must contain a half-integral number of wavelengths, or

$$d \sin \theta = (m + \tfrac{1}{2})\lambda \qquad m = 0, 1, 2, \ldots \text{ (minima).} \tag{43–2}$$

If a lens is used as in Fig. 43–6, it may seem that a phase difference should develop between the rays beyond the plane represented by S_2b, the path lengths between this plane and P being clearly different. In Section 42–5, however, we saw that for such parallel rays focused by a lens the *optical path lengths* are identical. Two rays with the same optical path lengths contain the same number of wavelengths, so that no phase difference will result because of the light passing through the lens.

▶ **Example 1.** The double-slit arrangement in Fig. 43-5 is illuminated with light from a mercury vapor lamp so filtered that only the strong green line ($\lambda = 5460$ A) is effective. The slits are 0.10 mm apart, and the screen on which the interference pattern appears is 20 cm away. What is the angular position of the first minimum? Of the tenth maximum?

At the first minimum we put $m = 0$ in Eq. 43-2, or

$$\sin\theta = \frac{(m + \frac{1}{2})\lambda}{d} = \frac{(\frac{1}{2})(546 \times 10^{-9}\text{ meter})}{0.10 \times 10^{-3}\text{ meter}} = 0.0027.$$

This value for $\sin\theta$ is so small that we can take it to be the value of θ, expressed in radians; expressed in degrees it is 0.16°.

At the tenth maximum (not counting the central maximum) we must put $m = 10$ in Eq. 43-1. Doing so and calculating as before leads to an angular position of 3.8°. For these conditions we see that the angular spread of the first dozen or so fringes is small.

Example 2. In Example 1 what is the linear distance on screen C between adjacent maxima?

If θ is small enough, we can use the approximation

$$\sin\theta \cong \tan\theta \cong \theta.$$

From Fig. 43-5 we see that

$$\tan\theta = \frac{y}{D}.$$

Substituting this into Eq. 43-1 for $\sin\theta$ leads to

$$y = m\frac{\lambda D}{d} \qquad m = 0, 1, 2, \ldots \text{ (maxima)}.$$

The positions of any two adjacent maxima are given by

$$y_m = m\frac{\lambda D}{d}$$

and

$$y_{m+1} = (m + 1)\frac{\lambda D}{d}.$$

Their separation Δy is found by subtracting:

$$\Delta y = y_{m+1} - y_m = \frac{\lambda D}{d}$$

$$= \frac{(546 \times 10^{-9}\text{ meter})(20 \times 10^{-2}\text{ meter})}{0.10 \times 10^{-3}\text{ meter}} = 1.09\text{ mm}.$$

As long as θ in Figs. 43-5 and 43-6 is small, the separation of the interference fringes is independent of m; that is, the fringes are evenly spaced. Note that if the incident light contains more than one wavelength the separate interference patterns, which will have different fringe spacings, will be superimposed. ◀

Equation 43-1 can be used to determine the wavelength of light; to quote Thomas Young:

> From a comparison of various experiments, it appears that the breadth of the undulations [that is, the wavelength] constituting the extreme red light must be supposed to be, in air, about one 36 thousandth of an inch, and those of the extreme violet about one 60 thousandth; the mean of the whole spectrum, with respect to the intensity of light, being one 45 thousandth.

Young's value for the average effective wavelength present in sunlight (1/45,000 in.) can be written as 5700 A, which agrees rather well with the wavelength at which the eye sensitivity is a maximum, 5550 A (see Fig. 40–2). It must not be supposed that Young's work was received without criticism. One of his contemporaries, evidently a firm believer in the corpuscular theory of light, wrote:

> We wish to raise our feeble voice against innovations that can have no other effect than to check the progress of science, and renew all those wild phantoms of the imagination which Bacon and Newton put to flight from her temple. This paper contains nothing that deserves the name of either experiment or discovery.

Needless to say, posterity has decided in favor of Young.

43–2 Coherence

Analysis of the derivation of Eqs. 43–1 and 43–2 shows that a fundamental requirement for the existence of well-defined interference fringes on screen C in Fig. 43–1 is that the light waves that travel from S_1 and S_2 to any point P on this screen must have a sharply defined phase difference ϕ that remains constant with time. If this condition is satisfied, a stable, well-defined fringe pattern will appear. At certain points P, ϕ will be given, independent of time, by $n\pi$ where $n = 1, 3, 5, \ldots$ so that the resultant intensity will be strictly zero and will remain so throughout the time of observation. At other points ϕ will be given by $n\pi$ where $n = 0, 2, 4 \ldots$ and the resultant intensity will be a maximum. Under these conditions the two beams emerging from slit S_1 and S_2 are said to be completely *coherent*.

Let the source in Fig. 43–1 be removed and let slits S_1 and S_2 be replaced by two completely independent light sources, such as two fine incandescent wires placed side by side in a glass envelope. No interference fringes will appear on screen C but only a relatively uniform illumination. We can interpret this if we make the reasonable assumption that for completely independent light sources the phase difference between the two beams arriving at P will vary with time in a random way. At a certain instant conditions may be right for cancellation and a short time later (perhaps 10^{-8} sec) they may be right for re-enforcement. This same random phase behavior holds for all points on screen C with the result that this screen is uniformly illuminated. The intensity at any point is equal to the sum of the intensities that each source S_1 and S_2 produces separately at that point. Under these conditions the two beams emerging from S_1 and S_2 are said to be completely *incoherent*.

Note that for completely coherent light beams one (1) combines the amplitudes vectorially, taking the (constant) phase difference properly into account, and then (2) squares this resultant amplitude to obtain a quantity proportional to the resultant intensity. For completely incoherent light beams, on the other hand, one (1) squares the individual amplitudes to obtain quantities proportional to the individual intensities and then (2) adds the individual intensities to obtain the resultant intensity. This procedure is in agreement with the experimental fact that for completely independent light sources the resultant intensity at every point is always greater than the intensity produced at that point by either light source acting alone.

It remains to investigate under what experimental conditions coherent or incoherent beams may be produced and to give an explanation for coherence in terms of the mode of production of the radiation. Consider first a parallel beam of microwave radiation emerging from an antenna connected by a coaxial cable to an oscillator based on an electromagnetic resonant cavity. The cavity oscillations (see Section 38-6) are completely periodic with time and produce, at the antenna, a completely periodic variation of **E** and **B** with time. The radiated wave at large enough distances from the antenna is well represented by Fig. 39-11. Note that (1) the wave has essentially infinite extent in time, including both future times ($t > 0$, say) and past times ($t < 0$); see Fig. 43-7*a*. At any point, as the wave passes by, the wave disturbance (i.e., **E** or **B**) varies with time in a perfectly periodic way. (2) The wavefronts at points far removed from the antenna are parallel planes of essentially infinite extent at right angles to the propagation direction. At any instant of time the wave disturbance varies with distance along the propagation direction in a perfectly periodic way.

Two beams generated from a single traveling wave like that of Fig. 39-11 will be completely coherent. One way to generate two such beams is to put an opaque screen containing two slits in the path of the beam. The waves emerging from the slits will always have a constant phase difference at any point in the region in which they overlap and interference fringes will be produced. Coherent radio beams can also be readily established, as can coherent elastic waves in solids, liquids and gases. The two prongs of the vibrating tapper in Fig. 43-4, for example, generate two coherent waves in the water of the ripple tank.

The technique of producing two beams from a single beam (and thus from a single source) tests specifically whether the wavefronts in the single parallel beam are truly planes, that is, whether all points in a plane at right angles to the direction of propagation have the same phase at any given instant. By dividing the beam in another way it is possible to test whether the beam is truly periodic over a large number of cycles of oscillation. This can be done, as we show in detail in Section 43-6, by inserting in the beam at 45° to it a thin sheet of material possessing the property that two beams are produced, one (which will be at right angles to the incident beam) by reflection and a second (which will be in the direction of the incident beam) by transmission. In the visible region such a sheet, called a *half-silvered mirror*, may be formed from a glass plate by depositing on it an appropriately thin film of silver. By appropriate use of mirrors (see Section 43-7) these two sub-beams can be recombined into a single beam traveling in a chosen direction. If the beams travel different distances before they are recombined, we are comparing in the combined beam a sample of the original beam with another sample an arbitrarily large number of cycles away. If the original beam is truly periodic in space and time, the two sub-beams will be completely coherent and interference fringes will be produced when they are recombined.

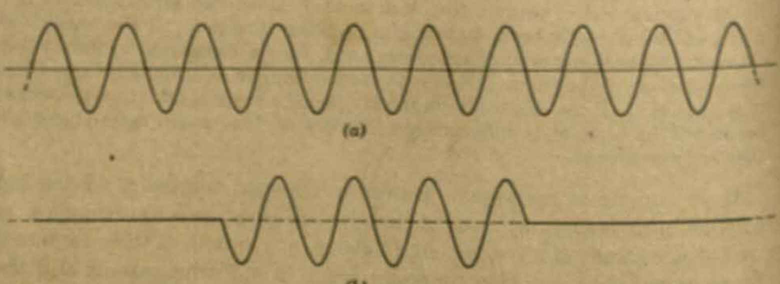
If we turn from microwave sources to common sources of visible light, such as incandescent wires or an electric discharge passing through a gas, we become aware of a fundamental difference. In both of these sources the fundamental light emission processes occur in individual atoms and these atoms do not act together in a cooperative (i.e., *coherent*) way. The act of light emission by a single atom takes, in a typical case, about 10^{-8} sec and the emitted light is properly described as a *wavetrain* (Fig. 43-7*b*) rather than as a wave (Fig. 43-7*a*). For emission times such as these the wavetrains are a few meters long.

Interference effects from ordinary light sources may be produced by putting a very narrow slit (S_0 in Fig. 43-1) directly in front of the source. This insures that the wavetrains that strike slits S_1 and S_2 in screen B in this figure originate from the same small region of the source. The diffracted beams emerging from S_1 and S_2 thus represent the same population of wavetrains and are coherent with respect to each other. If the phase of the light emitted from S_0 changes, this change is transmitted simultaneously to S_1 and S_2. Thus, at any point on screen C, a constant phase difference is maintained between the beams from these two slits and a stationary interference pattern occurs.

If the width of slit S_0 in Fig. 43-1 is gradually increased it will be observed experimentally that the maxima of the interference fringes become reduced in intensity and that the intensity in the fringe minima is no longer strictly zero. In other words, the fringes become less distinct. If S_0 is opened extremely wide, the lowering of the maximum intensity and the raising of the minimum intensity will be such that the fringes disappear, leaving only a uniform illumination. Under these conditions we say that the beams from S_1 and S_2 pass continuously from a condition of complete coherence to one of complete incoherence. When not at either of these two limits the beams are said to be partially coherent.

Partial coherence can also be demonstrated in two beams that are produced, as described on p. 1090 by inserting a "half-silvered mirror" in a beam at an angle of 45° to the direction of propagation. The two beams so produced, by reflection and transmission, traverse paths of different lengths before they are recombined. If the path difference is small compared to the average length of a wavetrain, the interference fringes will be sharply defined and will go essentially to zero at their minima. If the path difference is deliberately made longer, the fringes will become less distinct, and finally, when the path difference is larger than the average length of a wavetrain, the fringes will disappear altogether. Thus it is possible once again to progress smoothly in an experimental arrangement from complete coherence, through partial coherence, to complete incoherence.

The lack of coherence of the light from ordinary sources such as glowing wires is due to the fact that the emitting atoms do not act cooperatively


(*a*)

(*b*)

Fig. 43–7 (*a*) A section of an infinite wave and (*b*) a wavetrain.

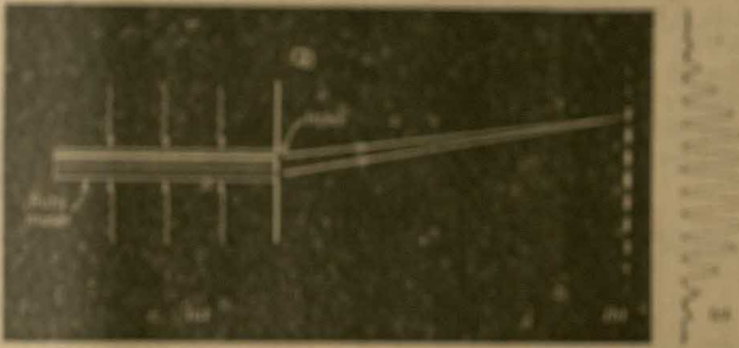


Fig. 43-6 A screen with two small holes is placed against the end of an optical maser. (a) The light passing through the holes forms an interference pattern on a strip of photographic film placed at (b). The fact that the pattern is formed shows that the light emitted from the maser is coherent across the beam cross section. At (b) is shown the image as formed on a photographic strip exposed in this manner. At (c) is an intensity plot of the film made by measuring the degree to which the film has been exposed. (Courtesy of D. F. Nelson and R. J. Collins, Bell Telephone Laboratories.)

(i.e., *coherently*). Since 1960 it has proved possible to construct sources of visible light in which the atoms *do* act cooperatively and in which the emitted light is highly coherent. Such devices are called *optical masers* or *lasers*;* their light output is extremely monochromatic, intense, and highly collimated. The coherence of the emitted light can be demonstrated by placing a screen with two holes in it in the beam emerging from the laser. An interference pattern results, as Fig. 43-6 shows. These methods permit, for the first time, a degree of control of visible light approaching that possible for radio and for microwaves. The practical applications of lasers, which include the amplification of weak light signals, the use of light beams as highly efficient carriers of information from point to point (see problem 12), and the production of high temperatures by intense local heating, remain to be fully exploited.

43-3 Intensity in Young's Experiment

Let us assume that the electric field components of the two waves in Fig. 43-5 vary with time at point P as

$$E_1 = E_0 \sin \omega t \tag{43-3}$$

and

$$E_2 = E_0 \sin (\omega t + \phi) \tag{43-4}$$

where ω $(= 2\pi\nu)$ is the angular frequency of the waves and ϕ is the phase difference between them. Note that ϕ depends on the location of point P,

---

[*] Laser, like maser, is a coined word, meaning "light amplification through stimulated emission of radiation." The interested student is referred to an article called "Optical Masers" by Arthur L. Schawlow, *Scientific American*, June 1961.

which, in turn, for a fixed geometrical arrangement, is described by the angle θ in Figs. 43–5 and 43–6. We assume that the slits are so narrow that the diffracted light from each slit illuminates the central portion of the screen uniformly. This means that near the center of the screen E_0 is independent of the position of P, that is, of the value of θ

The resultant wave disturbance * at P is found from

$$E = E_1 + E_2 \tag{43–5}$$

and (see Eq. 19–17) proves to be

$$E = E_\theta \sin (\omega t + \beta), \tag{43–6a}$$

where

$$\beta = \tfrac{1}{2}\phi \tag{43–6b}$$

and

$$E_\theta = 2E_0 \cos \beta = E_m \cos \beta. \tag{43–6c}$$

E_m, the maximum possible amplitude for E_θ, is equal to twice the amplitude of the combining waves ($= 2E_0$), corresponding to complete reinforcement. The student should verify Eq. 43–6 carefully. The amplitude E_θ of the resultant wave disturbance, which determines the intensity of the interference fringes, will turn out to depend strongly on the value of θ, that is, on the location of point P in Figs. 43–5 and 43–6.

In Section 19–6 we showed that the intensity of a wave I, measured perhaps in watts/meter^2, is proportional to the square of its amplitude. For the resultant wave then, ignoring the proportionality constant,

$$I_\theta \propto E_\theta^2. \tag{43–7}$$

This relationship seems reasonable if we recall (Eq. 30–27) that the energy density in an electric field is proportional to the *square* of the electric field strength. This is true for rapidly varying electric fields, such as those in a light wave, as well as for static fields.

The ratio of the intensities of two light waves is the ratio of the squares of the amplitudes of their electric fields. If I_θ is the intensity of the resultant wave at P and I_0 is the intensity that a single wave acting alone would produce, then

$$\frac{I_\theta}{I_0} = \left(\frac{E_\theta}{E_0}\right)^2 \tag{43–8}$$

Combining with Eq. 43–6c leads to

$$I_\theta = 4I_0 \cos^2 \beta = I_m \cos^2 \beta. \tag{43–9}$$

---

* The electric field **E** in the light wave rather than the magnetic field **B** is normally identified with the "wave disturbance" because the effects of **B** on the human eye and on various light detectors are exceedingly small. Radiation pressure (Section 40–2) is one such effect. Note too that, although Eq. 43–5 should be a vector equation, in most cases of interest the **E** vectors in the two interfering waves are closely parallel so that an algebraic equation suffices.

Note that the intensity of the resultant wave at any point P varies from zero [for a point at which $\phi\ (= 2\beta) = \pi$, say] to I_m, which is four times the intensity I_0 of each individual wave [for a point at which $\phi\ (= 2\beta) = 0$, say]. Let us compute I_θ as a function of the angle θ in Figs. 43-5 or 43-6.

The phase difference ϕ in Eq. 43-4 is associated with the path difference S_1b in Fig. 43-5 or 43-6. If S_1b is $\frac{1}{2}\lambda$, ϕ will be π; if S_1b is λ, ϕ will be 2π, etc. This suggests that

$$\frac{\text{phase difference}}{2\pi} = \frac{\text{path difference}}{\lambda},$$

$$\phi = \frac{2\pi}{\lambda}(d \sin\theta),$$

or, finally, from Eq. 43-6*b*,

$$\beta = \tfrac{1}{2}\phi = \frac{\pi d}{\lambda}\sin\theta. \qquad (43\text{–}10)$$

This expression for β can be substituted into Eq. 43-9 for I_θ, yielding the latter quantity as a function of θ. For convenience we collect here the expressions for the amplitude and the intensity in double-slit interference.

$$\left.\begin{array}{ll}\text{[Eq. 43–6c]} & E_\theta = E_m \cos\beta \\ \text{[Eq. 43–9]} & I_\theta = I_m \cos^2\beta \\ \text{[Eq. 43–10]} & \beta\ (= \tfrac{1}{2}\phi) = \dfrac{\pi d}{\lambda}\sin\theta \end{array}\right] \begin{array}{ll}\text{interference} & (43\text{–}11a) \\ \text{from narrow} & (43\text{–}11b) \\ \text{slits (that is,} & \\ a \ll \lambda) & (43\text{–}11c)\end{array}$$

To find the positions of the intensity maxima, we put

$$\beta = m\pi \qquad m = 0, 1, 2, \ldots$$

in Eq. 43-11*b*. From Eq. 43-11*c* this reduces to

$$d \sin\theta = m\lambda \qquad m = 0, 1, 2, \ldots \text{ (maxima)},$$

which is the equation derived in Section 43-1 (Eq. 43-1). To find the intensity minima we write

$$\frac{\pi d \sin\theta}{\lambda} = (m + \tfrac{1}{2})\pi \qquad m = 0, 1, 2, \ldots \text{ (minima)},$$

which reduces to the previously derived Eq. 43-2.

Figure 43-9 shows the intensity pattern for double-slit interference. The horizontal solid line is I_0; this describes the (uniform) intensity pattern on the screen if one of the slits is covered up. If the two sources were incoherent the intensity would be uniform over the screen and would be $2I_0$; see the horizontal dashed line in Fig. 43-9. For coherent sources we expect the

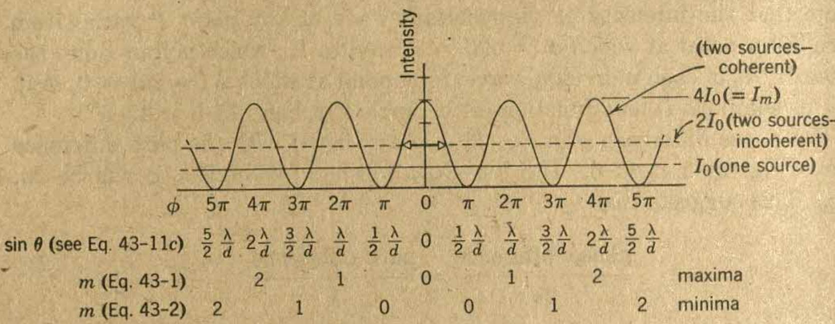


Fig. 43–9 The intensity pattern for double-slit interference. The heavy arrow in the central peak represents the half-width of the peak. This figure is constructed on the assumption that the two interfering waves each illuminate the central portion of the screen uniformly, that is, I_0 is independent of position as shown.

energy to be merely redistributed over the screen, since energy is neither created nor destroyed by the interference process. Thus the *average* intensity in the interference pattern should be $2I_0$, as for incoherent sources. This follows at once if, in Eq. 43–11*b*, we substitute one-half for the cosine-squared term and if we recall that $I_m = 4I_0$. We have seen several times that the average value of the square of a sine or a cosine term over one or more half-cycles is one-half.

43–4 Adding Wave Disturbances

In Section 43–3 we combined two time-varying wave disturbances, namely

$$E_1 = E_0 \sin \omega t \tag{43–3}$$

and

$$E_2 = E_0 \sin (\omega t + \phi), \tag{43–4}$$

which have the same angular frequency ω and amplitude E_0 but which have a phase difference ϕ between them. In this case the result (Eqs. 43–11*a* and *c*) is easily obtained algebraically.

In later chapters we will want to add larger numbers of wave disturbances, often an infinite number, with infinitesimal individual amplitudes. Since analytic methods become more difficult in such cases we describe a graphical method, illustrating it by rederiving Eq. 43–11*a*.

A sinusoidal wave disturbance such as that represented by Eq. 43–3 can be represented graphically, using a rotating vector. In Fig. 43–10*a* a vector of magnitude E_0 is allowed to rotate about the origin in a counterclockwise direction with an angular frequency ω. Following electrical engineering practice we call such a rotating vector a *phasor*. The alternating wave disturbance E_1 (Eq. 43–3) is represented by the projection of this phasor on the vertical axis.

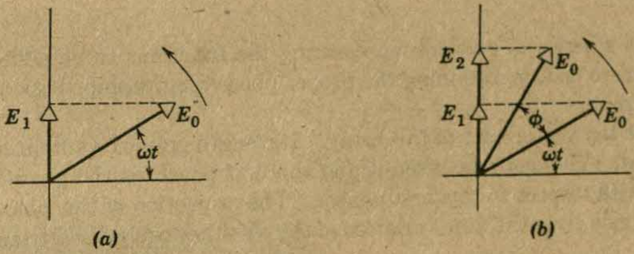
A second wave disturbance E_2, which has the same amplitude E_0 but a phase difference ϕ with respect to E_1,

$$E_2 = E_0 \sin (\omega t + \phi), \tag{43-4}$$

can be represented graphically (Fig. 43–10*b*) as the projection on the vertical axis of a second phasor of magnitude E_0 which makes an angle ϕ with the first phasor. As this figure shows, the sum E of E_1 and E_2 is the sum of the projections of the two phasors on the vertical axis. This is revealed more clearly if the phasors are redrawn, as in Fig. 43–10*c*, placing the foot of one arrow at the head of the other, maintaining the proper phase difference, and letting the whole assembly rotate counterclockwise about the origin.

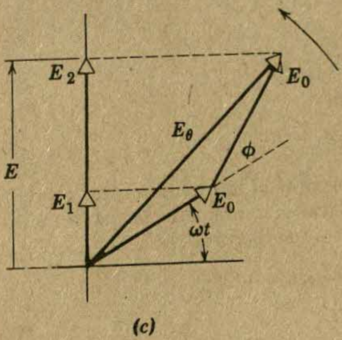
In Fig. 43–10*c* E can also be regarded as the projection on the vertical axis of a phasor of length E_θ, which is the vector sum of the two phasors of magnitude E_0. Note that the (algebraic) sum of the projections of the two phasors is equal to the projection of the (vector) sum of the two phasors.

In most problems in optics we are concerned only with the *amplitude* E_θ of the resultant wave disturbance and not with its time variation. This is because the eye and other common measuring instruments respond to the resultant intensity of the light (that is, to the square of the amplitude) and cannot detect the rapid time variations that characterize visible light. For sodium light, for example (λ = 5890 A), the frequency ν (= $\omega/2\pi$) is 5.1 × 10^{14} cycles/sec. Often, then, we need not consider the rotation of the phasors but can confine our attention to finding the magnitude of the resultant phasor.



(a)

(b)



(c)

Fig. 43–10 (*a*) A wave disturbance E_1 is represented by a rotating vector or *phasor*. (*b*) Two wave disturbances E_1 and E_2, with a phase difference ϕ between them, are so represented. These two phasors can represent the two wave disturbances in the double-slit problem; see Eqs. 43–3 and 43–4. (*c*) Another way of drawing (*b*).

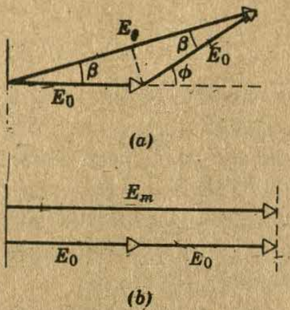


Fig. 43–11 (a) A construction to find the amplitude E_θ of two wave disturbances of amplitude E_0 and phase difference ϕ. (b) The maximum possible amplitude for these wave disturbances occurs for $\phi =$ (and has the value $E_m = 2E_0$.

Figure 43–11*a* shows the phasors for double-slit interference at time $t = 0$; compare Fig. 43–10*c*. We see that

$$E_\theta = 2E_0 \cos \beta = E_m \cos \beta,$$

in which, from the theorem that the exterior angle of a triangle (ϕ) is equal to the sum of its two opposite interior angles ($\beta + \beta$),

$$\beta = \tfrac{1}{2}\phi.$$

This is exactly the result arrived at earlier algebraically; compare Eqs. 43–11*a* and *c*.

In a more general case we might want to find the resultant of a number (>2) of sinusoidally varying wave disturbances. The general procedure is the following:

1. Construct a series of phasors representing the functions to be added. Draw them end to end, maintaining the proper phase relationships between adjacent phasors.
2. Construct the vector sum of this array. Its length gives the amplitude of the resultant. The angle between it and the first phasor is the phase of the resultant with respect to this first phasor. The projection of this phasor on the vertical axis gives the time variation of the resultant wave disturbance.

▶ **Example 3.** Find graphically the resultant $E(t)$ of the following wave disturbances:

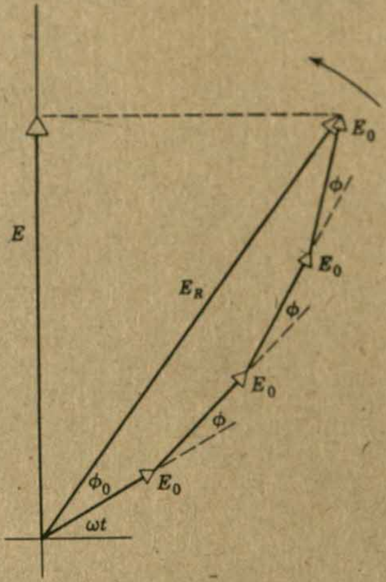
$$E_1 = 10 \sin \omega t$$
$$E_2 = 10 \sin (\omega t + 15°)$$
$$E_3 = 10 \sin (\omega t + 30°)$$
$$E_4 = 10 \sin (\omega t + 45°).$$

Figure 43–12 in which E_0 equals 10, shows the assembly of four phasors that represents these functions. Their vector sum, by graphical measurement, has an amplitude E_R of 38 and a phase ϕ_0 with respect to E_1 of 23°. In other words

$$E(t) = E_1 + E_2 + E_3 + E_4 = 38 \sin (\omega t + 23°).$$

Check this result by trigonometric calculation. ◀

Fig. 43-12 Example 3. Four wave disturbances are added graphically, using the method of phasors.



43-5 Interference from Thin Films

The colors of soap bubbles, oil slicks, and other thin films are the result of interference. Figure 43-13 shows interference effects in a thin vertical film of soapy water illuminated by monochromatic light.

Figure 43-14 shows a film of uniform thickness d and index of refraction n, the eye being focused on spot a. The film is illuminated by a broad source of monochromatic light S. There exists on this source a point P such that two rays, identified by the single and double arrows, respectively, can leave P and enter the eye as shown, after passing through point a. These two rays follow

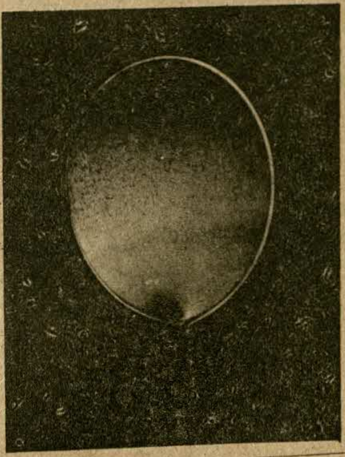


Fig. 43-13 A soapy water film on a wire loop viewed by reflected light. The black segment at the top is not a tear. It arises because the film, by drainage, is so thin here that there is destructive interference between the light reflected from its front surface and that reflected from its back surface. Recall that these two waves differ in phase by 180°

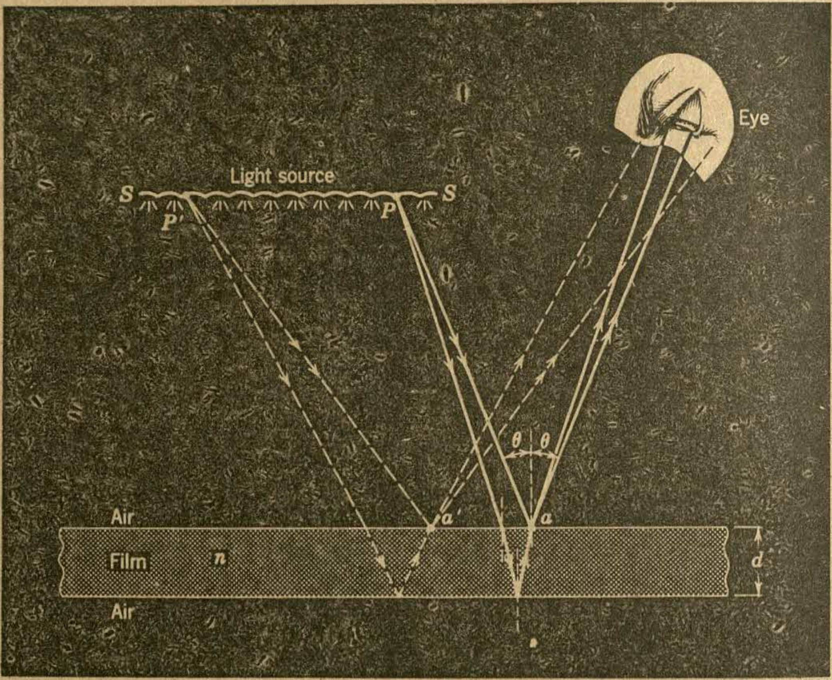


Fig. 43–14 Interference by reflection from a thin film, assuming an extended source S.

different paths in going from P to the eye, one being reflected from the upper surface of the film, the other from the lower surface. Whether point a appears bright or dark depends on the nature of the interference between the two waves that diverge from a. These waves are coherent because they both originate from the same point P on the light source.

If the eye looks at another part of the film, say a', the light that enters the eye must originate from a different point P' of the extended source, as suggested by the dashed lines in Fig. 43–14.

For near-normal incidence ($\theta \cong 0$ in Fig. 43–14) the geometrical path difference for the two rays from P will be close to $2d$. We might expect the resultant wave reflected from the film near a to be an interference maximum if the distance $2d$ is an integral number of wavelengths. This statement must be modified for two reasons.

First, the wavelength must refer to the wavelength of the light in the film λ_n and not to its wavelength in air λ; that is, we are concerned with optical path lengths rather than geometrical path lengths. The wavelengths λ and λ_n (see Eq. 41–11) are related by

$$\lambda_n = \lambda/n. \tag{43–12}$$

To bring out the second point, let us assume that the film is so thin that $2d$ is very much less than a wavelength. The phase difference between the two waves would be close to zero on our assumption, and we would expect such a film to appear bright on reflection. However, it appears dark. This is clear from Fig. 43-13, in which the action of gravity produces a wedge-shaped film, extremely thin at its top edge. As drainage continues, the dark area increases in size. To explain this and many similar phenomena, we assume that one or the other of the two rays of Fig. 43-14 suffers an abrupt phase change of π (= 180°) associated either with reflection at the air-film interface or transmission through it. As it turns out, the ray reflected from the upper surface suffers this phase change. The other ray is not changed abruptly in phase, either on transmission through the upper surface or on reflection at the lower surface.

In Section 19-9 we discussed phase changes on reflection for transverse waves in strings. To extend these ideas, consider the composite string of Fig. 43-15, which consists of two parts with different masses per unit length, stretched to a given tension. If a pulse moves to the right in Fig. 43-15a, approaching the junction, there will be a reflected and a transmitted pulse, the reflected pulse being *in phase* with the incident pulse. In Fig. 43-15b the situation is reversed, the incident pulse now being in the less massive string. In this case the reflected pulse will differ in phase from the incident pulse by π (= 180°). In each case the transmitted pulse will be in phase with the incident pulse.

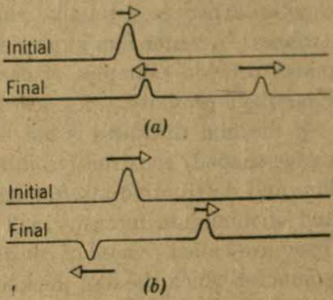


Fig. 43-15 Phase changes on reflection at a junction between two stretched composite strings. (*a*) Incident pulse in heavy string and (*b*) incident pulse in light string.

Figure 43-15*a* suggests a light wave in glass, say, approaching a surface beyond which there is a less optically dense medium (one of lower index of refraction) such as air. Figure 43-15*b* suggests a light wave in air approaching glass. To sum up the optical situation, when reflection occurs from an interface beyond which the medium has a *lower* index of refraction, the reflected wave undergoes *no phase change;* when the medium beyond the interface has a *higher* index, there is a phase change of π.* The transmitted wave does not experience a change of phase in either case.

---

* These statements, which can be proved rigorously from Maxwell's equations (see also Section 43-6), must be modified for light falling on a less dense medium at an angle such that total internal reflection occurs. They must also be modified for reflection from metallic surfaces.

We are now able to take into account both factors that determine the nature of the interference, namely, differences in optical path length and phase changes on reflection. For the two rays of Fig. 43–14 to combine to give a *maximum* intensity, assuming normal incidence, we must have

$$2d = (m + \tfrac{1}{2})\lambda_n \qquad m = 0, 1, 2, \ldots.$$

The term $\frac{1}{2}\lambda_n$ is introduced because of the phase change on reflection, a phase change of 180° being equivalent to half a wavelength. Substituting λ/n for λ_n yields finally

$$2dn = (m + \tfrac{1}{2})\lambda \qquad m = 0, 1, 2, \ldots \text{ (maxima)}. \tag{43–13}$$

The condition for a *minimum* intensity is

$$2dn = m\lambda \qquad m = 0, 1, 2, \ldots \text{ (minima)}. \tag{43–14}$$

These equations hold when the index of refraction of the film is either greater or less than the indices of the media on *each* side of the film. Only in these cases will there be a relative phase change of 180° for reflections at the two surfaces. A water film in air and an air film in the space between two glass plates provide examples of cases to which Eqs. 43–13 and 43–14 apply. Example 5 provides a case in which they do not apply.

If the film thickness is not uniform, as in Fig. 43–13, where the film is wedge-shaped, constructive interference will occur in certain parts of the film and destructive interference will occur in others. Lines of maximum and of minimum intensity will appear—these are the interference fringes. They are called *fringes of constant thickness*, each fringe being the locus of points for which the film thickness d is a constant. If the film is illuminated with white light rather than monochromatic light, the light reflected from various parts of the film will be modified by the various constructive or destructive interferences that occur. This accounts for the brilliant colors of soap bubbles and oil slicks.

Only if the film is "thin," which implies that d is no more than a few wavelengths of light, will fringes of the type described, that is, fringes that appear localized on the film and associated with a variable film thickness, be possible. For very thick films (say $d \cong 1$ cm), the path difference between the two rays of Fig. 43–14 will be many wavelengths and the phase difference at a given point on the film will change rapidly as we move even a small distance away from a. For "thin" films, however, the phase difference at a also holds for reasonably nearby points; there is a characteristic "patch brightness" for any point on the film, as Fig. 43–13 shows. Interference fringes can be produced for thick films; they are not localized on the film but are at infinity. See Section 43–7.

▶ **Example 4.** A water film ($n = 1.33$) in air is 3200 A thick. If it is illuminated with white light at normal incidence, what color will it appear to be in reflected light?

By solving Eq. 43–13 for λ,

$$\lambda = \frac{2dn}{m + \frac{1}{2}} = \frac{(2)(3200\text{ A})(1.33)}{m + \frac{1}{2}} = \frac{8500\text{ A}}{m + \frac{1}{2}} \qquad \text{(maxima)}.$$

From Eq. 43-14 the minima are given by

$$\lambda = \frac{8500\text{ A}}{m} \qquad \text{(minima)}.$$

Maxima and minima occur for the following wavelengths:

m	0 (max)	1 (min)	1 (max)	2 (min)	2 (max)
λ, A	17000	8500	5700	4250	3400

Only the maximum corresponding to $m = 1$ lies in the visible region (see Fig. 40-2); light of this wavelength appears yellow-green. If white light is used to illuminate the film, the yellow-green component will be enhanced when viewed by reflection.

Example 5. *Nonreflecting glass.* Lenses are often coated with thin films of transparent substances like MgF_2 ($n = 1.38$) in order to reduce the reflection from the glass surface, using interference. How thick a coating is needed to produce a minimum reflection at the center of the visible spectrum (5500 A)?

We assume that the light strikes the lens at near-normal incidence (θ is exaggerated for clarity in Fig. 43-16), and we seek destructive interference between rays r and r_1. Equation 43-14 does not apply because in this case a phase change of 180° is associated with *each* ray, for at *both* the upper and lower surfaces of the MgF_2 film the reflection is from a medium of greater index of refraction.

There is no net change in phase produced by the two reflections, which means that the optical path difference for destructive interference is $(m + \frac{1}{2})\lambda$ (compare Eq. 43-13), leading to

$$2dn = (m + \tfrac{1}{2})\lambda \qquad m = 0, 1, 2, \ldots \text{ (minima)}.$$

Solving for d and putting $m = 0$ yields

$$d = \frac{(m + \frac{1}{2})\lambda}{2n} = \frac{\lambda}{4n} = \frac{5500\text{ A}}{(4)(1.38)} = 1000\text{ A}.$$

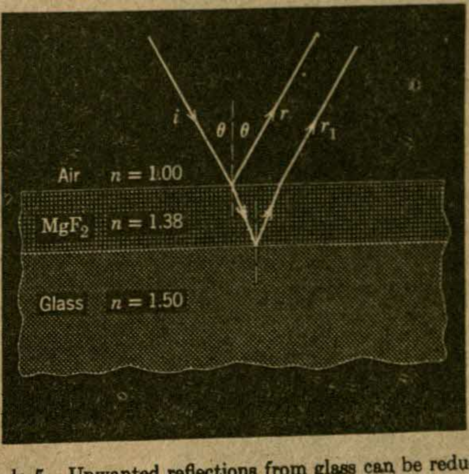


Fig. 43-16 Example 5. Unwanted reflections from glass can be reduced by coating the glass with a thin transparent film.

Example 6. *Newton's rings.* Figure 43–17 shows a lens of radius of curvature R resting on an accurately plane glass plate and illuminated from above by light of wavelength λ. Figure 43–18 shows that circular interference fringes (Newton's rings) appear, associated with the variable thickness air film between the lens and the plate. Find the radii of the circular interference maxima.

Here it is the ray from the *bottom* of the (air) film rather than from the top that undergoes a phase change of 180°, for it is the one reflected from a medium of higher refractive index. The condition for a maximum remains unchanged, however (Eq. 43–13), and is

$$2d = (m + \tfrac{1}{2})\lambda \qquad m = 0, 1, 2, \ldots, \tag{43-15}$$

the index of refraction of the air film being assumed to be unity. From Fig. 43–17 we can write

$$d = R - \sqrt{R^2 - r^2} = R - R\left[1 - \left(\frac{r}{R}\right)^2\right]^{1/2}$$

If $r/R \ll 1$, the square bracket can be expanded by the binomial theorem, keeping only two terms, or

$$d = R - R\left[1 - \frac{1}{2}\left(\frac{r}{R}\right)^2 + \cdots\right] \cong \frac{r^2}{2R}.$$

Combining with Eq. 43–15 yields

$$r = \sqrt{(m + \tfrac{1}{2})\lambda R} \qquad m = 0, 1, 2, \ldots,$$

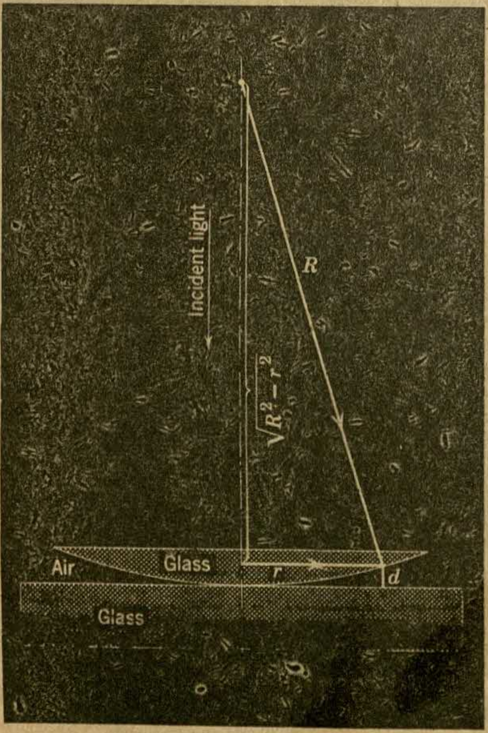


Fig. 43–17 Example 6. Apparatus for observing Newton's rings.

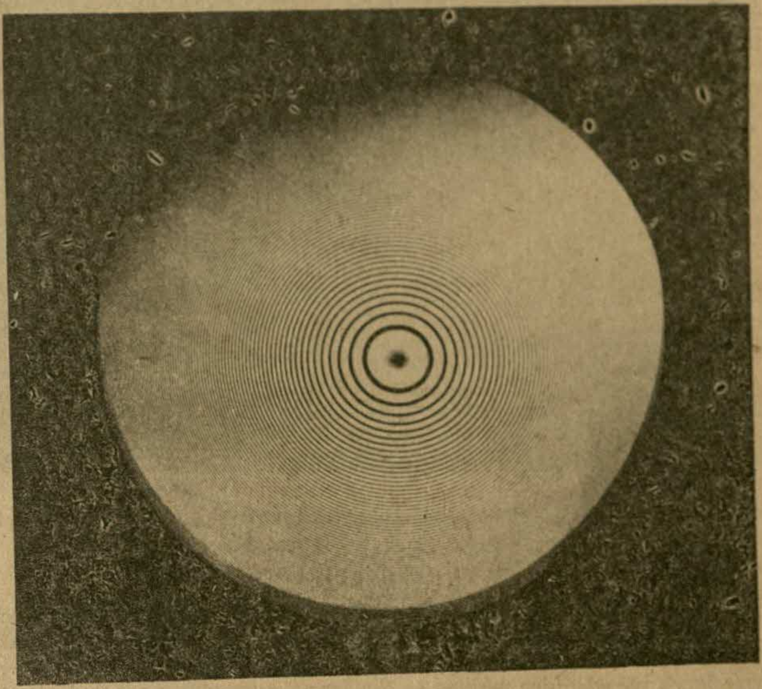


Fig. 43–18 Example 6. Newton's rings. (Courtesy of Bausch and Lomb Optical Co.)

which gives the radii of the bright rings. If white light is used, each spectrum component will produce its own set of circular fringes, the sets all overlapping. ◀

43–6 Phase Changes on Reflection

G. G. Stokes (1819–1903) used the principle of optical reversibility to investigate the reflection of light at an interface between two media. The principle states that if there is no absorption of light a light ray that is reflected or refracted will retrace its original path if its direction is reversed. This reminds us that any mechanical system can run backward as well as forward, provided there is no absorption of energy because of friction, etc.

Figure 43–19*a* shows a wave of amplitude E reflected and refracted at a surface separating media 1 and 2, where $n_2 > n_1$. The amplitude of the reflected wave is $r_{12}E$, in which r_{12} is an *amplitude reflection coefficient*. The amplitude of the refracted wave is $t_{12}E$, where t_{12} is an *amplitude transmission coefficient*.

We consider only the possibility of phase changes of 0 or 180°. If $r_{12} = +0.5$, for example, we have a reduction in amplitude on reflection by one-half and no change in phase. For $r_{12} = -0.5$ we have a phase change of 180° because

$$E \sin(\omega t + 180°) = -E \sin \omega t.$$

Figure 43–19*b* suggests that if we reverse these two rays they should combine to produce the original ray reversed in direction. Ray $r_{12}E$, identified by the single arrows in the figure, is reflected and refracted, producing the rays of amplitudes

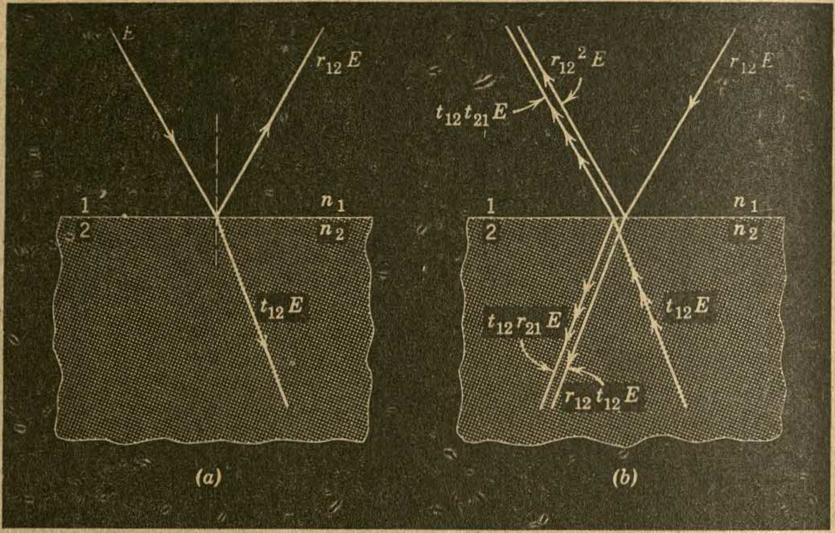


Fig. 43–19 (*a*) A ray is reflected and refracted at an air-glass interface. (*b*) The optically reversed situation; the two rays in the lower left must cancel.

r_{12}^2E and $r_{12}t_{12}E$. Ray $t_{12}E$, identified by the triple arrows, is also reflected and refracted, producing the rays of amplitudes $t_{12}t_{21}E$ and $t_{12}r_{21}E$ as shown. Note that r_{12} describes a ray in medium 1 reflected from medium 2, and r_{21} describes a ray in medium 2 reflected from medium 1. Similarly, t_{12} describes a ray that passes from medium 1 to medium 2; t_{21} describes a ray that passes from medium 2 to medium 1.

The two rays in the upper left of Fig. 43–19*b* must be equivalent to the incident ray of Fig. 43–19*a*, reversed; the two rays in the lower left of Fig. 43–19*b* must cancel. This second requirement leads to

$$r_{12}t_{12}E + t_{12}r_{21}E = 0,$$

or

$$r_{12} = -r_{21}.$$

This result tells us that if we compare a wave reflected from medium one with one reflected from medium 2, they behave differently in that one or the other undergoes a phase change of 180°. We must rely on experiment to show that, as we pointed out earlier, the ray reflected from the more optically dense medium is the one that experiences the phase change of 180°.

43–7 Michelson's Interferometer

An interferometer is a device that can be used to measure lengths or changes in length with great accuracy by means of interference fringes. We describe the form originally built by Michelson in 1881.

Consider light that leaves point P on extended source S (Fig. 43–20) and falls on half-silvered mirror M. This mirror has a silver coating just thick enough to transmit half the incident light and to reflect half; in the figure we have assumed that this mirror, for convenience, possesses negligible thickness. At M the light divides into two waves. One proceeds by transmission toward mirror M_1; the other proceeds by reflection toward M_2. The waves are reflected at each of these mirrors and are sent back along their directions of incidence, each wave eventually entering the eye E. Since the waves are coherent, being derived from the same point on the source, they will interfere.

If the mirrors M_1 and M_2 are exactly perpendicular to each other, the effect is that of light from an extended source S falling on a uniformly thick slab of air, between glass, whose thickness is equal to $d_2 - d_1$. Interference fringes appear, caused by small changes in the angle of incidence of the light from different points on the extended source as it strikes the equivalent air film. For *thick* films a path difference of one wavelength can be brought about by a very small change in the angle of incidence.

If M_2 is moved backward or forward, the effect is to change the thickness of the equivalent air film. Suppose that the center of the (circular) fringe pattern appears bright and that M_2 is moved just enough to cause the first bright circular fringe to move to the center of the pattern. The path of the light beam striking M_2 has been changed by one wavelength. This means (because the light passes twice through the equivalent air film) that the mirror must have moved one-half a wavelength.

The interferometer is used to measure changes in length by counting the number of interference fringes that pass the field of view as mirror M_2 is moved. Length measurements made in this way can be accurate if large numbers of fringes are counted.

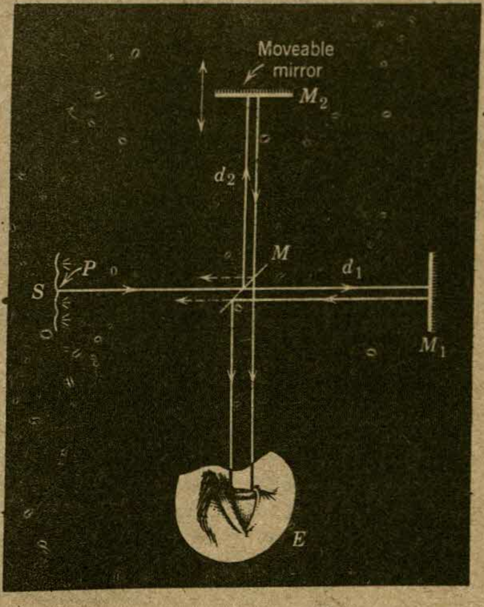


Fig. 43-20 Michelson's interferometer, showing the path of a particular ray originating at point P of an extended source S.

Michelson measured the length of the standard meter, kept in Paris, in terms of the wavelength of certain monochromatic red light emitted from a light source containing cadmium. He showed that the standard meter was equivalent to 1,553,163.5 wavelengths of the red cadmium light.

Physicists have long speculated on the advantages of discarding the standard meter bar as the basic standard of length and of *defining* the meter in terms of the wavelength of some carefully chosen monochromatic radiation. This would make the primary length standard readily available in laboratories all over the world. It would improve the accuracy of length measurements, since one would no longer need to compare an unknown object with a standard object (the meter bar), using interferometer techniques, but could measure the unknown object *directly* and in an absolute sense, using these techniques. There is the additional advantage that if the standard meter bar were destroyed it could never be replaced, whereas light sources and interferometers will (presumably) always be available.

In 1961 such an atomic standard of length was adopted by international agreement. Quoting from an article * describing the event:

> The wavelength of the orange-red light of krypton-86 has replaced the platinum iridium bar as the world standard of length. Formerly the wavelength of this light was defined as a function of the length of the meter bar. Now the meter is defined as a multiple (1,650,763.73) of the wavelength of the light.

The light from krypton-86 was used in preference to that from cadmium or other sources because it produces sharper interference fringes in the interferometer over the long optical paths sometimes used in length measurement.

43–8 Michelson's Interferometer and Light Propagation

In Section 40–4 we presented Einstein's hypothesis, now well verified, that in free space light is propagated with the same speed c no matter what the relative velocity of the source and the observer may be. We pointed out that this hypothesis contradicted the views of nineteenth-century physicists regarding wave propagation. It was difficult for these physicists, trained as they were in the classical physics of the time, to believe that a wave could be propagated without a medium. If such a medium could be established, the speed c of light would naturally be construed as the speed *with respect to that medium*, just as the speed of sound always refers to a medium such as air.

Although no medium for light propagation was obvious, the physicists postulated one, called the *ether* † and hypothesized that its properties were such that it was undetectable by ordinary means such as weighing.

In 1881 (24 years before Einstein's hypothesis) A. A. Michelson set himself the task of forcing the ether, assuming that it existed, to submit to direct physical verification. In particular, Michelson, later joined by E. W. Morley, tried to measure the speed u with which the earth moves through the ether. Michelson's interferometer was their instrument of choice for this now-famous Michelson-Morley experiment.

The earth together with the interferometer moving with velocity $\mathbf{u}$ through the ether is equivalent to the interferometer at rest with the ether streaming through it with velocity $-\mathbf{u}$, as shown in Fig. 43–21. Consider a wave moving along the path MM_1M and one moving along MM_2M. The first corresponds classically to a man

---

* *Scientific American*, p.75, December 1960.

† More fully, the *luminiferous* (or light-carrying) *ether*.

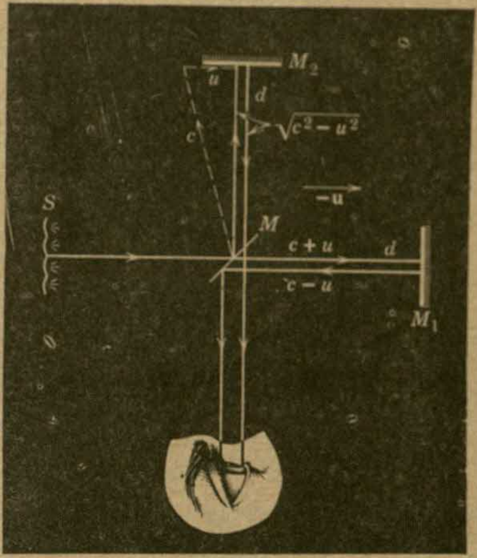


Fig. 43-21 The "ether" is streaming with velocity $-\mathbf{u}$ through Michelson's interferometer. The wave speeds shown are on the basis of the (incorrect) ether hypothesis.

rowing a boat a distance d downstream and the same distance upstream; the second corresponds to rowing a boat a distance d across a stream and back.

On the ether hypothesis the speed of light on the path MM_1 is $c + u$; on the return path M_1M it is $c - u$. The time required for the complete trip is

$$t_1 = \frac{d}{c+u} + \frac{d}{c-u} = d\frac{2c}{c^2 - u^2} = \frac{2d}{c}\frac{1}{1 - (u/c)^2}.$$

The speed of light, *on the ether hypothesis*, for path MM_2 is $\sqrt{c^2 - u^2}$, as Fig. 43-21 suggests. This same speed holds for the return path M_2M, so that the time required for this complete path is

$$t_2 = \frac{2d}{\sqrt{c^2 - u^2}} = \frac{2d}{c}\frac{1}{\sqrt{1 - (u/c)^2}}.$$

The difference of time for the two paths is

$$\begin{aligned}\Delta t &= t_1 - t_2\\ &= \frac{2d}{c}\left\{\left[1 - \left(\frac{u}{c}\right)^2\right]^{-1} - \left[1 - \left(\frac{u}{c}\right)^2\right]^{-1/2}\right\}.\end{aligned}$$

Assuming $u/c \ll 1$, we can expand the quantities in the square brackets by using the binomial theorem, retaining only the first two terms. This leads to

$$\begin{aligned}\Delta t &= \frac{2d}{c}\left\{\left[1 + \left(\frac{u}{c}\right)^2 + \cdots\right] - \left[1 + \frac{1}{2}\left(\frac{u}{c}\right)^2 + \cdots\right]\right\}\\ &= \frac{2d}{c}\left\{\frac{1}{2}\left(\frac{u}{c}\right)^2\right\} = \frac{du^2}{c^3}. \qquad (43\text{-}16)\end{aligned}$$

Now let the entire interferometer be rotated through 90°. This will interchange the roles of the two light paths, MM_1M now being the "cross-stream" path and MM_2M the "down- and upstream" path. The time difference between the two waves entering the eye is also reversed; this changes the phase difference between the combining waves and alters the positions of the interference maxima. *The experiment consists of looking for a shift of the interference fringes as the apparatus is rotated.*

The *change* in time difference is $2\Delta t$, which corresponds to a fringe shift of $2\Delta t/T$ where $T\,(=\lambda/c)$ is the period of vibration of the light. The expected maximum shift in the number of fringes on a 90° rotation (see Eq. 43–16) is

$$\Delta N = \frac{2\Delta t}{T} = \frac{2\Delta tc}{\lambda} = \frac{2d}{\lambda}\left(\frac{u}{c}\right)^2. \tag{43–17}$$

In the Michelson-Morley interferometer let $d = 11$ meters (obtained by multiple eflection in the interferometer) and $\lambda = 5.9 \times 10^{-7}$ meter. If u is assumed to be oughly the orbital speed of the earth, then $u/c \cong 10^{-4}$. The expected maximum ringe shift when the interferometer is rotated through 90° is then

$$\Delta N = \frac{2d}{\lambda}\left(\frac{u}{c}\right)^2 = \frac{(2)(11\text{ meters})}{5.9\times 10^{-7}\text{ meter}}(10^{-4})^2 = 0.4.$$

Even though a shift of only about 0.4 of a fringe was expected, Michelson and Morley were confident that they could observe a shift of 0.01 fringe. *They found from their experiment, however, that there was no observable fringe shift!*

The analogy between a light wave in the supposed ether and a boat moving in water, which seemed so evident in 1881, is simply incorrect. The derivation based on this analogy is incorrect for light waves. When the analysis is carried through on Einstein's hypothesis, the observed negative result is clearly predicted, the speed of light being c for all paths. The motion of the earth around the sun and the rotation of the interferometer have, in Einstein's view, no effect whatever on the speed of the light waves in the interferometer.

It should be made clear that although Einstein's hypothesis is completely consistent with the negative result of the Michelson-Morley experiment this experiment standing alone cannot serve as a proof for Einstein's hypothesis. Einstein said that no number of experiments, however large, could prove him right but that a single experiment could prove him wrong. Our present-day belief in Einstein's hypothesis rests on consistent agreement in a large number of experiments designed to test it. The "single experiment" that might prove Einstein wrong has never been found.

QUESTIONS

1. Is Young's experiment an interference experiment or a diffraction experiment, or both?

2. Do interference effects occur for sound waves? Recall that sound is a longitudinal wave and that light is a transverse wave.

3. In Young's double-slit interference experiment, using a monochromatic laboratory light source, why is screen A in Fig. 43-1 necessary? What would happen if one gradually enlarged the hole in this screen?

4. Describe the pattern of light intensity on screen C in Fig. 43-5 if one slit is covered with a red filter and the other with a blue filter, the incident light being white.

5. Is coherence important in reflection and refraction?

6. Define carefully, and distinguish between, the angles θ and ϕ that appear in Eq. 43-10.

7. If one slit in Fig. 43-5 is covered, what change occurs in the intensity of light in the center of the screen?

8. What changes occur in the pattern of interference fringes if the apparatus of Fig. 43-5 is placed under water?

9. What are the requirements for a maximum intensity when viewing a thin film by *transmitted* light?

10. In a Newton's rings experiment, is the central spot, as seen by reflection, dark or light? Explain.

11. Why must the film of Fig. 43-14 be "thin" for us to see an interference pattern of the type described?

12. Why do coated lenses (see Example 5) look purple by reflected light?

13. A person wets his eyeglasses to clean them. As the water evaporates he notices that for a short time the glasses become markedly more nonreflecting. Explain.

14. A lens is coated to reduce reflection, as in Example 5. What happens to the energy that had previously been reflected? Is it absorbed by the coating?

15. Very small changes in the angle of incidence do not change the interference conditions much for "thin" films but they do change them for "thick" films. Why?

16. The directional characteristics of a certain radar antenna as a receiver of radiation are known. What can be said about its directional characteristics as a transmitter?

17. A person in a dark room, looking through a small window, can see a second person standing outside in bright sunlight. The second person cannot see the first person. Is this a failure of the principle of optical reversibility? Assume no absorption of light.

18. Why is it necessary to rotate the interferometer in the Michelson-Morley experiment?

19. How is the negative result of the Michelson-Morley experiment interpreted according to Einstein's theory of relativity?

20. If interference between light waves of different frequencies is possible, one should observe light beats, just as one obtains sound beats from two sources of sound with slightly different frequencies. Discuss how one might experimentally look for this possibility.

21. In Young's double-slit experiment suppose that screen A in Fig. 43-1 contained *two* very narrow parallel slits instead of one. (a) Show that if the spacing between these slits is properly chosen the interference fringes can be made to disappear. (b) Under these conditions, would you call the beams emerging from slits S_1 and S_2 in screen B coherent? They do not produce interference fringes. (c) Discuss what would happen to the interference fringes in the case of a single slit in screen A if the slit width were gradually increased.

PROBLEMS

1. Design a double-slit arrangement that will produce interference fringes 1° apart on a distant screen. Assume sodium light ($\lambda = 5890$ A).

2. Two point sources in Fig. 43–22 emit coherent waves. Show that curves, such as that given, over which the phase difference for rays r_1 and r_2 is a constant are hyperbolas. Extend the analysis to three dimensions. (Hint: A constant phase difference implies a constant difference in length between r_1 and r_2.)

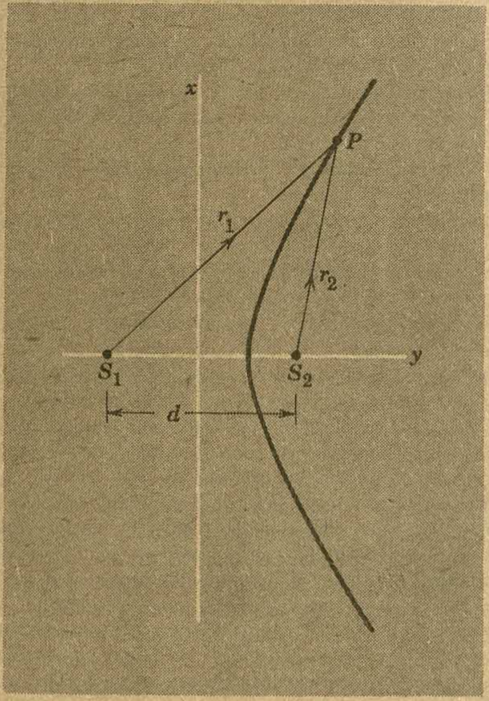


Fig. 43–22

3. A double-slit arrangement produces interference fringes for sodium light ($\lambda = 5890$ A) that are 0.20° apart. What is the angular fringe separation if the entire arrangement is immersed in water?

4. A double-slit arrangement produces interference fringes for sodium light ($\lambda = 5890$ A) that are 0.20° apart. For what wavelength would the angular separation be 10% greater?

5. A thin flake of mica ($n = 1.58$) is used to cover one slit of a double-slit arrangement. The central point on the screen is occupied by what used to be the seventh bright fringe. If $\lambda = 5500$ A, what is the thickness of the mica?

6. Sodium light ($\lambda = 5890$ A) falls on a double slit of separation $d = 2.0$ mm. D in Fig. 43–5 is 4 cm. What per cent error is made in locating the tenth bright fringe if it is *not* assumed that $D \gg d$?

7. Sodium light ($\lambda = 5890$ A) falls on a double slit of separation $d = 0.20$ mm. A thin lens ($f = +1.0$ meter) is placed near the slit as in Fig. 43–6. What is the linear fringe separation on a screen placed in the focal plane of the lens?

8. S_1 and S_2 in Fig. 43-23 are effective point sources of radiation, excited by the same oscillator. They are coherent and in phase with each other. Placed 4.0 meters apart, they emit equal amounts of power in the form of 1.0-meter wavelength electromagnetic waves. (*a*) Find the positions of the first (that is, the nearest), the second, and the third maxima of the received signal, as the detector is moved out along Ox. (*b*) Is the intensity at the nearest minimum equal to zero? Justify your answer.

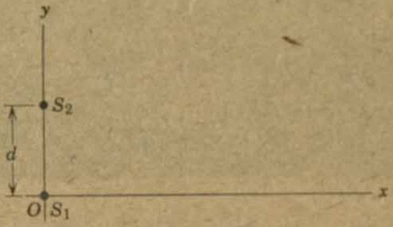


Fig. 43-23

9. One of the slits of a double-slit system is wider than the other, so that the amplitude of the light reaching the central part of the screen from one slit, acting alone, is twice that from the other slit, acting alone. Derive an expression for I_θ in terms of θ, corresponding to Eqs. 43-11*b* and *c*.

10. Show that the half-width $\Delta\theta$ of the double-slit interference fringes (see arrow in Fig. 43-9) is given by

$$\Delta\theta = \frac{\lambda}{2d}$$

if θ is small enough so that $\sin\theta \cong \theta$.

11. Find the sum of the following quantities (*a*) by the vector method and (*b*) analytically:

$$y_1 = 10 \sin \omega t$$

$$y_2 = 8 \sin (\omega t + 30°).$$

12. Add the following quantities graphically, using the vector method:

$$y_1 = 10 \sin \omega t$$

$$y_2 = 15 \sin (\omega t + 30°)$$

$$y_3 = 5 \sin (\omega t - 45°).$$

13. The frequency width of a standard television channel is 4×10^6 cycles/sec. If it were possible to transmit television signals in the visible portion of the electromagnetic spectrum, possibly by using optical maser or laser techniques, how many such standard channels would be available for allocation in the visible spectrum? Assume that this spectrum extends from 4000 to 7000 A.

14. In a Newton's rings experiment the radius of curvature R of the lens is 5.0 meters and its diameter is 2.0 cm. (*a*) How many rings are produced? (*b*) How many rings would be seen if the arrangement were immersed in water ($n = 1.33$)? Assume that $\lambda = 5890$ A.

15. The diameter of the tenth bright ring in a Newton's rings apparatus changes from 1.40 to 1.27 cm as a liquid is introduced between the lens and the plate. Find the index of refraction of the liquid.

16. A broad source of light ($\lambda = 6800$ A) illuminates normally two glass plates 12 cm long that touch at one end and are separated by a wire 0.048 mm in diameter at the other (Fig. 43-24). How many bright fringes appear over the 12-cm distance?

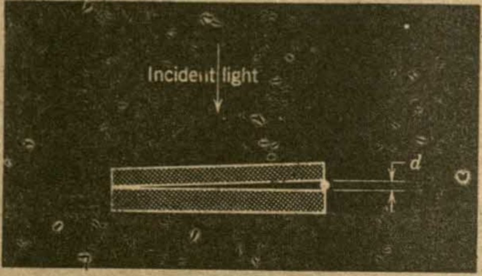


Fig. 43-24

17. A thin film 4×10^{-5} cm thick is illuminated by white light normal to its surface. Its index of refraction is 1.5. What wavelengths within the visible spectrum will be intensified in the reflected beam?

18. White light reflected at perpendicular incidence from a soap film has, in the visible spectrum, an interference maximum at 6000 A and a minimum at 4500 A, with no minimum in between. If $n = 1.33$ for the film, what is the film thickness, assumed uniform?

19. A plane wave of monochromatic light falls normally on a uniformly thin film of oil which covers a glass plate. The wavelength of the source can be varied continuously. Complete destructive interference of the reflected light is observed for wavelengths of 5000 and 7000 A and for no other wavelengths in between. If the index of refraction of the oil is 1.30 and that of the glass is 1.50, find the thickness of the oil film.

20. In Example 5 assume that there is zero reflection for light of wavelength 5500 A at normal incidence. Calculate the factor by which the reflection is diminished by the coating at 4500 and at 6500 A.

21. A plane monochromatic light wave in air falls at normal incidence on a thin film of oil which covers a glass plate. The wavelength of the source may be varied continuously. Complete destructive interference in the reflected beam is observed for wavelengths of 5000 and 7000 A and for no other wavelength in between. The index of refraction of glass is 1.50. Show that the index of refraction of the oil must be less than 1.50.

22. White light reflected at perpendicular incidence from a soap bubble has, in the visible spectrum, a single interference maximum (at $\lambda = 6000$ A) and a single minimum at the violet end of the spectrum. If $n = 1.33$ for the film, calculate its thickness.

23. If mirror M_2 in Michelson's interferometer is moved through 0.233 mm, 792 fringes are counted. What is the wavelength of the light?

24. A thin film with $n = 1.40$ for light of wavelength 5890 A is placed in one arm of a Michelson interferometer. If a shift of 7.0 fringes occurs, what is the film thickness?

25. (*a*) What is the wavelength of the orange-red line of krypton-86 in angstrom units? Take 1 A $= 10^{-10}$ meter exactly. (*b*) Does this question really make sense, in view of the fact that the meter is defined in terms of this wavelength? Explain.

26. A Michelson interferometer is used with a sodium discharge tube as a light source. The yellow sodium light consists of two wavelengths, 5890 and 5896 A. It is observed that the interference pattern disappears and reappears periodically as one moves mirror M_2 in Fig. 43-20. (*a*) Explain this effect. (*b*) Calculate the change in path difference between two successive reappearances of the interference pattern.

fraction

CHAPTER 44

44-1 Introduction

Diffraction, which is illustrated in Fig. 42-3, is the bending of light around an obstacle such as the edge of a slit. We can see the diffraction of light by looking through a crack between two fingers at a distant light source such as a tubular neon sign or by looking at a street light through a cloth umbrella. Usually diffraction effects are small and must be looked for carefully. Also, most sources of light have an extended area so that a diffraction pattern produced by one point of the source will overlap that produced by another. Finally, common sources of light are not monochromatic. The patterns for the various wavelengths overlap and again the effect is less apparent.

Diffraction was discovered by Francesco Maria Grimaldi (1618–1663), and the phenomenon was known both to Huygens (1629–1695) and to Newton (1642–1727). Newton did not see in it any justification for a wave theory for light. Huygens, although he believed in a wave theory, did not believe in diffraction! He imagined his secondary wavelets to be effective only at the point of tangency to their common envelope, thus denying the possibility of diffraction. In his words:

> And thus we see the reasons why light . . . proceeds only in straight lines in such a way that it does not illuminate any object except when the path from the source to the object is open along such a line.

Fresnel (1788–1827) correctly applied Huygens' principle (which is called the Huygens-Fresnel principle in Europe) to explain diffraction. In these early days the light waves were believed to be mechanical waves in an all-

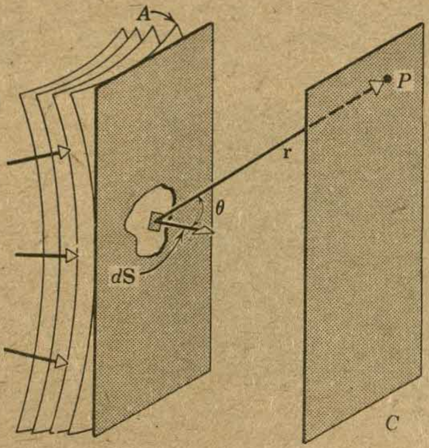


Fig. 44–1 Light is diffracted at the aperture in screen B and illuminates screen C. The intensity at P is found by dividing the wavefront at B into elementary radiators $d\mathbf{S}$ and combining their effects at P.

pervading ether. We have seen (Section 39–5) how Maxwell (1831–1879) showed that light waves were not mechanical in nature but electromagnetic. Einstein (1879–1955) rounded out our modern view of light waves by eliminating the need to postulate an ether (see Section 40–4).

Figure 44–1 shows a general diffraction situation. Surface A is a wavefront that falls on B, which is an opaque screen containing an aperture of arbitrary shape; C is a diffusing screen that receives the light that passes through this aperture. This pattern of light intensity on C can be calculated by subdividing the wavefront into elementary areas $d\mathbf{S}$, each of which becomes a source of an expanding Huygens' wavelet. The light intensity at an arbitrary point P is found by superimposing the wave disturbances (that is, the $\mathbf{E}$ vectors) caused by the wavelets reaching P from all these elementary radiators.

The wave disturbances reaching P differ in amplitude and in phase because (*a*) the elementary radiators are at varying distances from P, (*b*) the light leaves the radiators at various angles to the normal to the wavefront (see p. 1020), and (*c*) some radiators are blocked by screen B; others are not. Diffraction calculations—simple in principle—may become difficult in practice. The calculation must be repeated for every point on screen C at which we wish to know the light intensity. We followed exactly this program in calculating the double-slit intensity pattern in Section 43–3. The calculation there was simple because we assumed only two elementary radiators, the two narrow slits.

Figure 44–2*a* shows the general case of *Fresnel diffraction*, in which the light source and/or the screen on which the diffraction pattern is displayed are a finite distance from the diffracting aperture; the wavefronts that fall on the diffracting aperture in this case and that leave it to illuminate any point P of the diffusing screen are not planes; the corresponding rays are not parallel.

A simplification results if source S and screen C are moved to a large distance from the diffraction aperture, as in Fig. 44-2*b*. This limiting case is called *Fraunhofer diffraction*. The wavefronts arriving at the diffracting aperture from the distant source S are planes, and the rays associated with these wavefronts are parallel to each other. Similarly, the wavefronts arriving at any point P on the distant screen C are planes, the corresponding rays

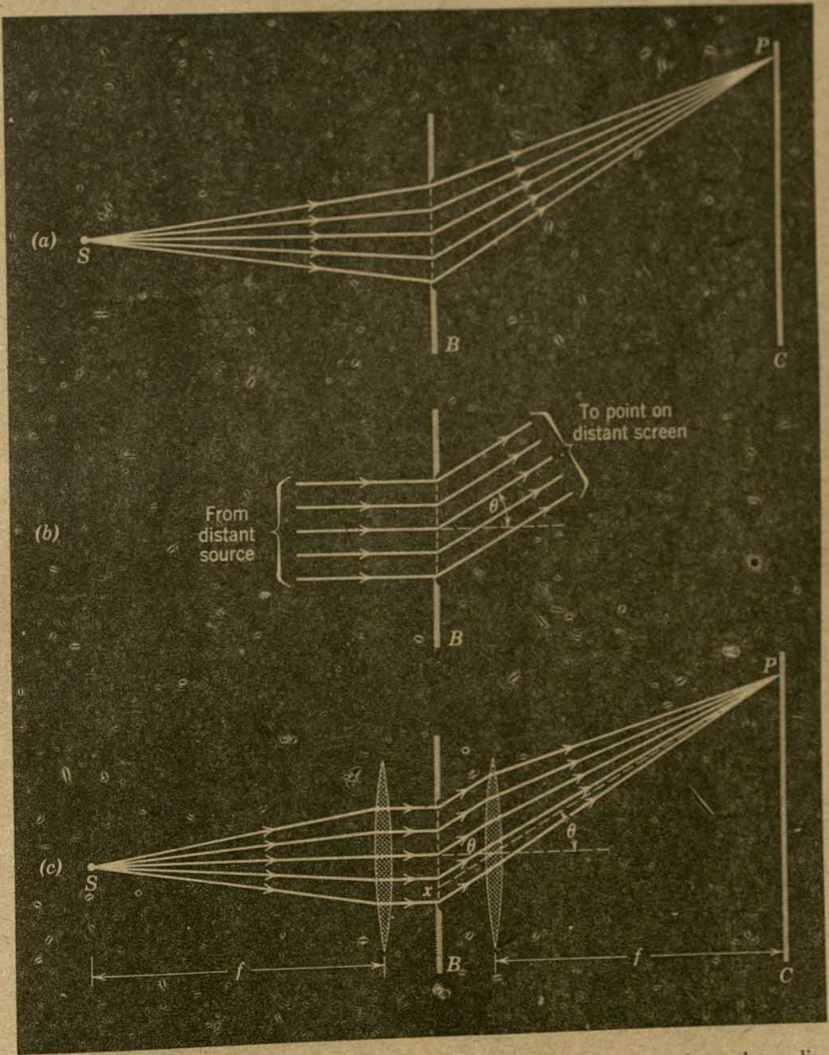


Fig. 44-2 (*a*) Fresnel diffraction. (*b*) Source S and screen C are moved to a large distance, resulting in Fraunhofer diffraction. (*c*) Fraunhofer diffraction conditions produced by lenses, leaving source S and screen C in their original positions.

also being parallel. Fraunhofer conditions can be established in the laboratory by using two converging lenses, as in Fig. 44-2*c*. The first of these converts the diverging wave from the source into a plane wave. The second lens causes plane waves leaving the diffracting aperture to converge to point *P*. All rays that illuminate *P* will leave the diffracting aperture parallel to the dashed line *Px* drawn from *P* through the center of this second (thin) lens. We assumed Fraunhofer conditions for Young's double-slit experiment in Section 43-1 (see Fig. 43-5).

Although Fraunhofer diffraction is a limiting case of the more general Fresnel diffraction, it is an important limiting case and is easier to handle mathematically. This book deals only with Fraunhofer diffraction.

44-2 Single Slit

Figure 44-3 shows a plane wave falling at normal incidence on a long narrow slit of width a. Let us focus our attention on the central point P_0 of screen C. The rays extending from the slit to P_0 all have the same optical path lengths, as we saw in Section 42-5. Since they are in phase at the plane of the slit, they will still be in phase at P_0, and the central point of the diffraction pattern that appears on screen C has a maximum intensity.

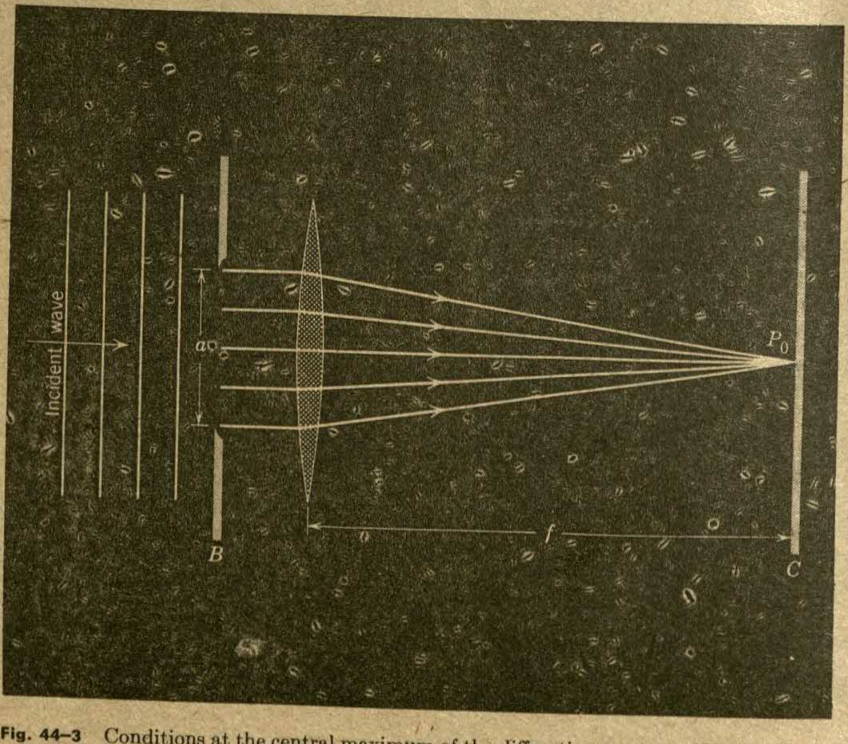


Fig. 44-3 Conditions at the central maximum of the diffraction pattern. The slit extends a distance above and below the figure, this distance being much greater than the slit width a.

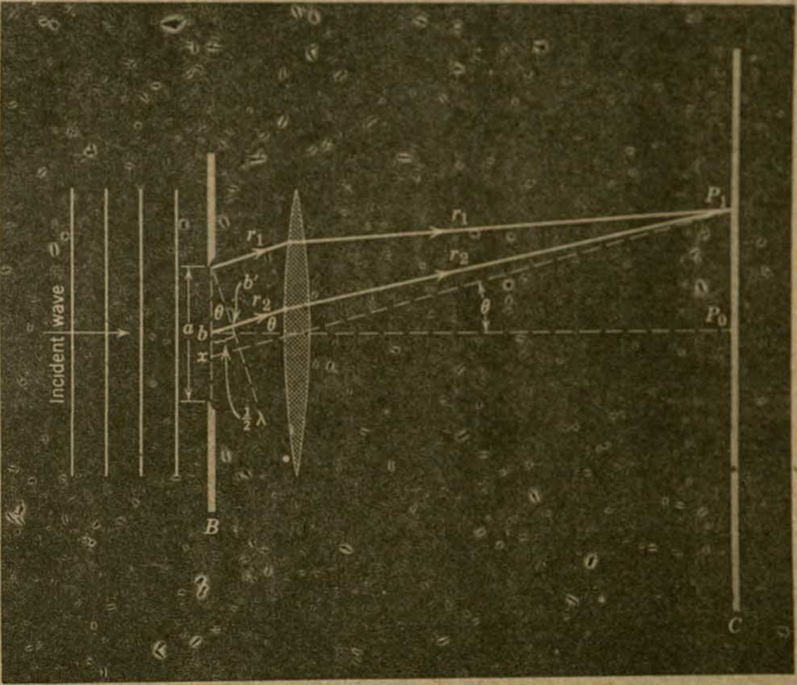


Fig. 44–4 Conditions at the first minimum of the diffraction pattern.

We now consider another point on the screen. Light rays which reach P_1 in Fig. 44–4 leave the slit at an angle θ as shown. Ray r_1 originates at the top of the slit and ray r_2 at its center. If θ is chosen so that the distance bb' in the figure is one-half a wavelength, r_1 and r_2 will be out of phase and will produce no effect at P_1.* In fact, every ray from the upper half of the slit will be canceled by a ray from the lower half, originating at a point $a/2$ below the first ray. The point P_1, the first minimum of the diffraction pattern, will have zero intensity (compare Fig. 42–3).

The condition shown in Fig. 44–4 is

$$\frac{a}{2}\sin\theta = \frac{\lambda}{2},$$

or

$$a\sin\theta = \lambda. \tag{44-1}$$

As we stated earlier (see Fig. 42–1), the central maximum becomes wider as the slit is made narrower. If the slit width is as small as one wavelength ($a = \lambda$), the first minimum occurs at $\theta = 90°$, which implies that the central maximum fills the entire forward hemisphere. We assumed a condition ap-

---

* Whatever phase relation exists between r_1 and r_2 at the plane represented by the sloping dashed line in Fig. 44–4 that passes through b' also exists at P_1, not being affected by the lens (see Section 42–5).

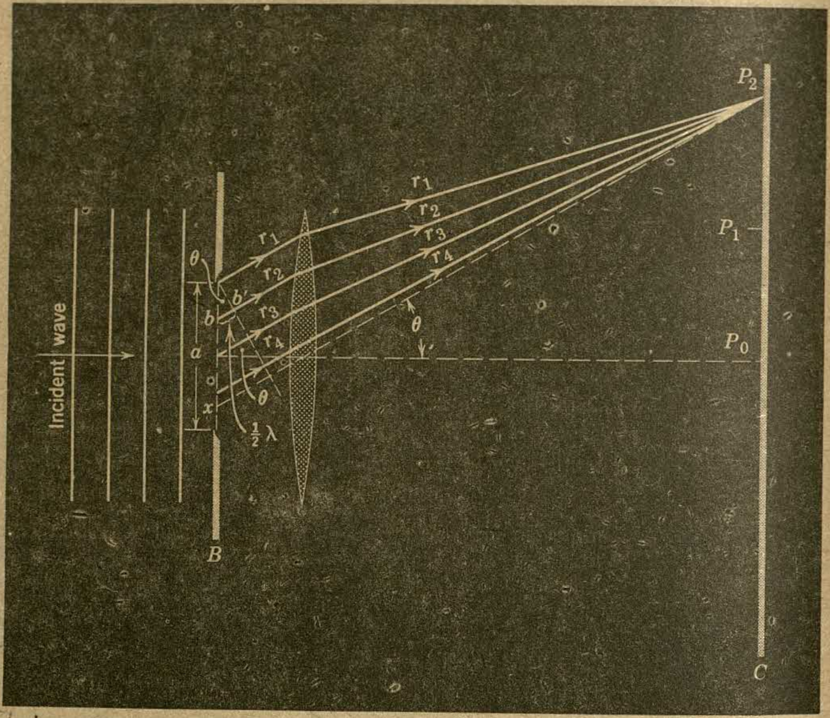


Fig. 44–5 Conditions at the second minimum of the diffraction pattern.

proaching this in our discussion of Young's double-slit interference experiment in Section 43–1.

In Fig. 44–5 the slit is divided into four equal zones, with a ray leaving the top of each zone. Let θ be chosen so that the distance bb' is one-half a wavelength. Rays r_1 and r_2 will then cancel at P_2. Rays r_3 and r_4 will also be half a wavelength out of phase and will also cancel. Consider four other rays, emerging from the slit a given distance below the four rays above. The two rays below r_1 and r_2 will cancel uniquely, as will the two rays below r_3 and r_4. We can proceed across the entire slit and conclude again that no light reaches P_2; we have located a second point of zero intensity.

The condition described (see Fig. 44–5) requires that

$$\frac{a}{4}\sin\theta = \frac{\lambda}{2},$$

or

$$a\sin\theta = 2\lambda.$$

By extension, the general formula for the minima in the diffraction pattern on screen C is

$$a\sin\theta = m\lambda \qquad m = 1, 2, 3, \ldots \text{ (minima).} \tag{44–2}$$

There is a maximum approximately halfway between each adjacent pair of minima. The student should consider the simplification that has resulted in this analysis by examining Fraunhofer (Fig. 44-2*c*) rather than Fresnel conditions (Fig. 44-2*a*).

▶ **Example 1.** A slit of width a is illuminated by white light. For what value of a will the first minimum for red light ($\lambda = 6500$ A) fall at $\theta = 30°$?

At the first minimum we put $m = 1$ in Eq. 44-2. Doing so and solving for a yields

$$a = \frac{m\lambda}{\sin\theta} = \frac{(1)(6500\text{ A})}{\sin 30°} = 13{,}000\text{ A}.$$

Note that the slit width must be twice the wavelength in this case.

Example 2. In Example 1 what is the wavelength λ' of the light whose first diffraction maximum (not counting the central maximum) falls at $\theta = 30°$, thus coinciding with the first minimum for red light?

This maximum is about halfway between the first and second minima. It can be found without too much error by putting $m = 1.5$ in Eq. 44-2, or

$$a \sin\theta \cong 1.5\lambda'.$$

From Example 1, however,
$$a\sin\theta = \lambda.$$

Dividing gives
$$\lambda' = \frac{\lambda}{1.5} = \frac{6500\text{ A}}{1.5} = 4300\text{ A}.$$

Light of this color is violet. The second maximum for light of wavelength 4300 A will *always* coincide with the first minimum for light of wavelength 6500 A, no matter what the slit width. If the slit is relatively narrow, the angle θ at which this overlap occurs will be relatively large. ◀

44-3 Single Slit—Qualitative

Figure 44-6 shows a slit of width a divided into N parallel strips of width Δx. Each strip acts as a radiator of Huygens' wavelets and produces a characteristic wave disturbance at point P, whose position on the screen, for a particular arrangement of apparatus, can be described by the angle θ.

If the strips are narrow enough—which we assume—all points on a given strip have essentially the same optical path length to P, and therefore all the light from the strip will have the same phase when it arrives at P. The amplitudes ΔE_0 of the wave disturbances at P from the various strips may be taken as equal if θ in Fig. 44-6 is not too large.

We limit our considerations to points that lie in, or infinitely close to, the plane of Fig. 44-6. It can be shown that this procedure is valid for a slit whose length is much greater than its width a. We made this same assumption tacitly both earlier in this chapter and in Chapter 43; see Figs. 43-5 and 44-3, for example.

The wave disturbances from adjacent strips have a constant phase difference $\Delta\phi$ between them at P given by

$$\frac{\text{phase difference}}{2\pi} = \frac{\text{path difference}}{\lambda},$$

or

$$\Delta\phi = \left(\frac{2\pi}{\lambda}\right)(\Delta x \sin\theta),$$

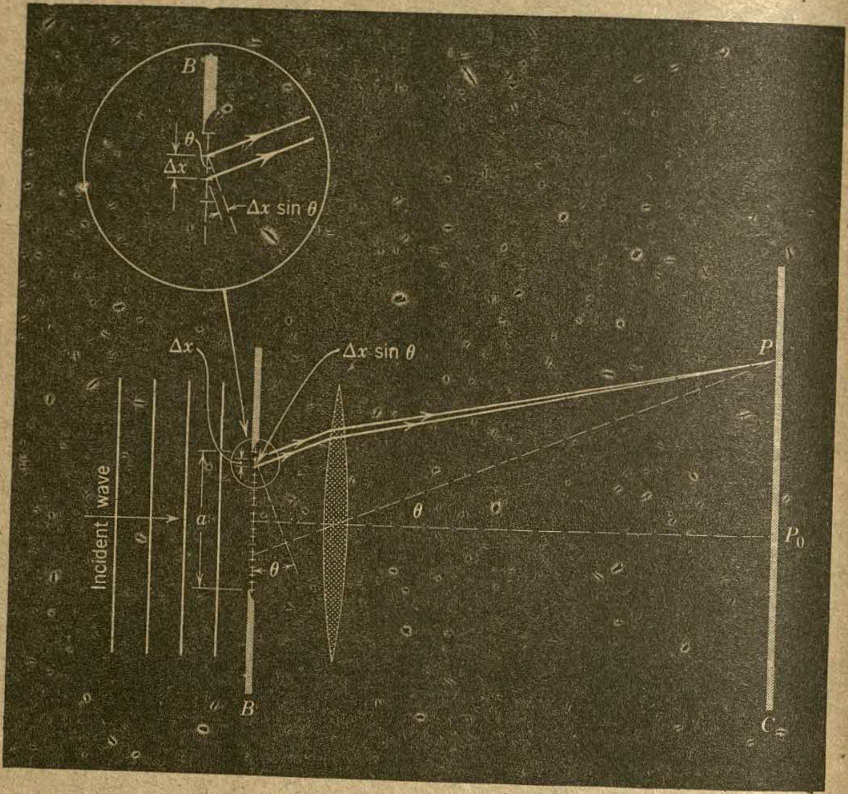


Fig. 44–6 A slit of width a is divided into N strips of width Δx. The insert shows conditions at the second strip more clearly. In the differential limit the slit is divided into an infinite number of strips (that is, $N \to \infty$) of differential width dx. For clarity in this and the following figure, we take $N = 18$.

where $\Delta x \sin \theta$ is, as the figure insert shows, the path difference for rays originating at the top edges of adjacent strips. Thus, at P, N vectors with the same amplitude ΔE_0, the same frequency, and the same phase difference $\Delta\phi$ between adjacent members combine to produce a resultant disturbance. We ask, for various values of $\Delta\phi$ [that is, for various points P on the screen, corresponding to various values of θ (see Eq. 44–3)], what is the amplitude E_θ of the resultant wave disturbance? We find the answer by representing the individual wave disturbances ΔE_0 by phasors and calculating the resultant phasor amplitude, as described in Section 43–4.

At the center of the diffraction pattern θ equals zero, and the phase shift between adjacent strips (see Eq. 44–3) is also zero. As Fig. 44–7a shows, the phasor arrows in this case are laid end to end and the amplitude of the resultant has its maximum value E_m. This corresponds to the center of the central maximum.

As we move to a value of θ other than zero, $\Delta\phi$ assumes a definite nonzero value (again see Eq. 44-3), and the array of arrows is now as shown in Fig. 44-7*b*. The resultant amplitude E_θ is less than before. Note that the length of the "arc" of small arrows is the same for both figures and indeed for all figures of this series. As θ increases further, a situation is reached (Fig. 44-7*c*) in which the chain of arrows curls around through 360°, the tip of the last arrow touching the foot of the first arrow. This corresponds to $E_\theta = 0$, that is, to the first minimum. For this condition the ray from the top of the slit (1 in Fig. 44-7*c*) is 180° out of phase with the ray from the center of the slit ($\frac{1}{2}N$ in Fig. 44-7*c*). These phase relations are consistent with Fig. 44-4, which also represents the first minimum.

As θ increases further, the phase shift continues to increase, and the chain of arrows coils around through an angular distance greater than 360°, as in Fig. 44-7*d*, which corresponds to the first maximum beyond the central maximum. This maximum is much smaller than the central maximum. In making this comparison, recall that the arrows marked E_θ in Fig. 44-7 correspond to the *amplitudes* of the wave disturbance and not to the in-

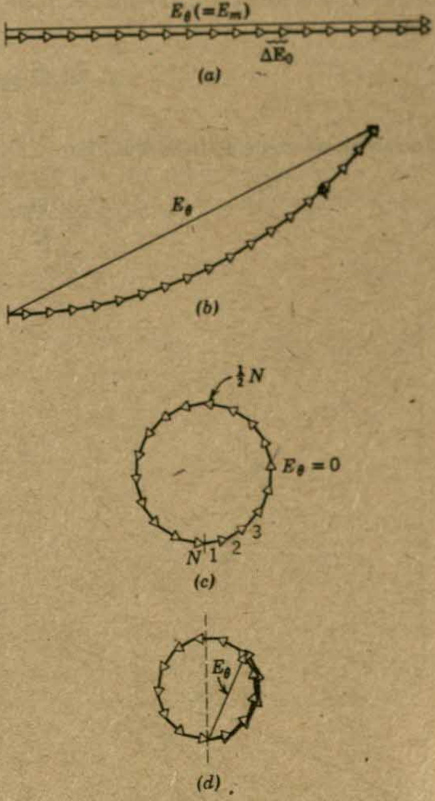


Fig. 44-7 Conditions at (*a*) the central maximum, (*b*) a direction slightly removed from the central maximum, (*c*) the first minimum, and (*d*) the first maximum beyond the central maximum for single-slit diffraction. This figure corresponds to $N = 18$ in Fig. 44-6.

tensity. The amplitudes must be squared to obtain the corresponding relative intensities (see Eq. 43–7).

44–4 Single Slit—Quantitative

The "arc" of small arrows in Fig. 44–8 shows the phasors representing, in amplitude and phase, the wave disturbances that reach an arbitrary point P on the screen of Fig. 44–6, corresponding to a particular angle θ. The resultant amplitude at P is E_θ. If we divide the slit of Fig. 44–6 into infinitesimal strips of width dx, the arc of arrows in Fig. 44–8 approaches the arc of a circle, its radius R being indicated in that figure. The length of the arc is E_m, the amplitude at the center of the diffraction pattern, for at the center of the pattern the wave disturbances are all in phase and this "arc" becomes a straight line as in Fig. 44–7*a*.

The angle ϕ in the lower part of Fig. 44–8 is revealed as the difference in phase between the infinitesimal vectors at the left and right ends of the arc E_m. This means that ϕ is the phase difference between rays from the top and the bottom of the slit of Fig. 44–6. From geometry we see that ϕ is also the angle between the two radii marked R in Fig. 44–8. From this figure we can write

$$E_\theta = 2R \sin \frac{\phi}{2}.$$

In radian measure ϕ, from the figure, is

$$\phi = \frac{E_m}{R}.$$

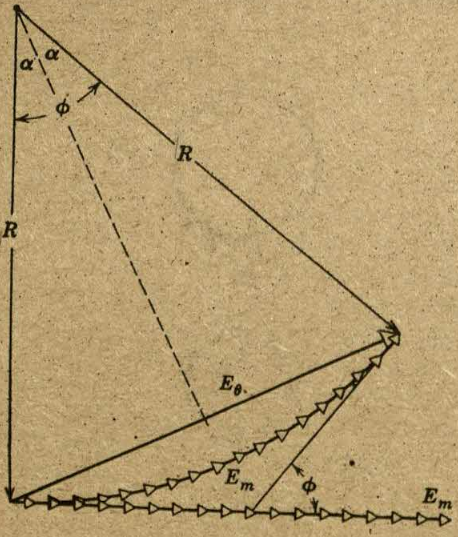


Fig. 44–8 A construction used to calculate the intensity in single-slit diffraction. The situation corresponds to that of Fig. 44–7.

Combining yields

$$E_\theta = \frac{E_m}{\phi/2}\sin\frac{\phi}{2},$$

or

$$E_\theta = E_m\frac{\sin\alpha}{\alpha}, \tag{44-4}$$

in which

$$\alpha = \frac{\phi}{2}. \tag{44-5}$$

From Fig. 44-6, recalling that ϕ is the phase difference between rays from the top and the bottom of the slit and that the path difference for these rays is $a\sin\theta$, we have

$$\frac{\text{phase difference}}{2\pi} = \frac{\text{path difference}}{\lambda},$$

or

$$\phi = \left(\frac{2\pi}{\lambda}\right)(a\sin\theta).$$

Combining with Eq. 44-5 yields

$$\alpha = \frac{\phi}{2} = \frac{\pi a}{\lambda}\sin\theta. \tag{44-6}$$

Equation 44-4, taken together with the definition of Eq. 44-6, gives the amplitude of the wave disturbance for a single-slit diffraction pattern at any angle θ. The intensity I_θ for the pattern is proportional to the square of the amplitude, or

$$I_\theta = I_m\left(\frac{\sin\alpha}{\alpha}\right)^2. \tag{44-7}$$

For convenience we display together, and renumber, the formulas for the amplitude and the intensity in single-slit diffraction.

$$\left.\begin{aligned} &[\text{Eq. } 44\text{-}4] & E_\theta &= E_m\frac{\sin\alpha}{\alpha} & &(44\text{-}8a)\\ &[\text{Eq. } 44\text{-}7] & I_\theta &= I_m\left(\frac{\sin\alpha}{\alpha}\right)^2 & &(44\text{-}8b)\\ &[\text{Eq. } 44\text{-}6] & \alpha\left(=\tfrac{1}{2}\phi\right) &= \frac{\pi a}{\lambda}\sin\theta & &(44\text{-}8c)\end{aligned}\right\}\ \text{single-slit diffraction}$$

Figure 44-9 shows plots of I_θ for several values of the ratio a/λ. Note that the pattern becomes narrower as a/λ is increased; compare this figure with Figs. 42-1 and 42-3.

Minima occur in Eq. 44-8b when

$$\alpha = m\pi \qquad m = 1, 2, 3, \ldots. \tag{44-9}$$

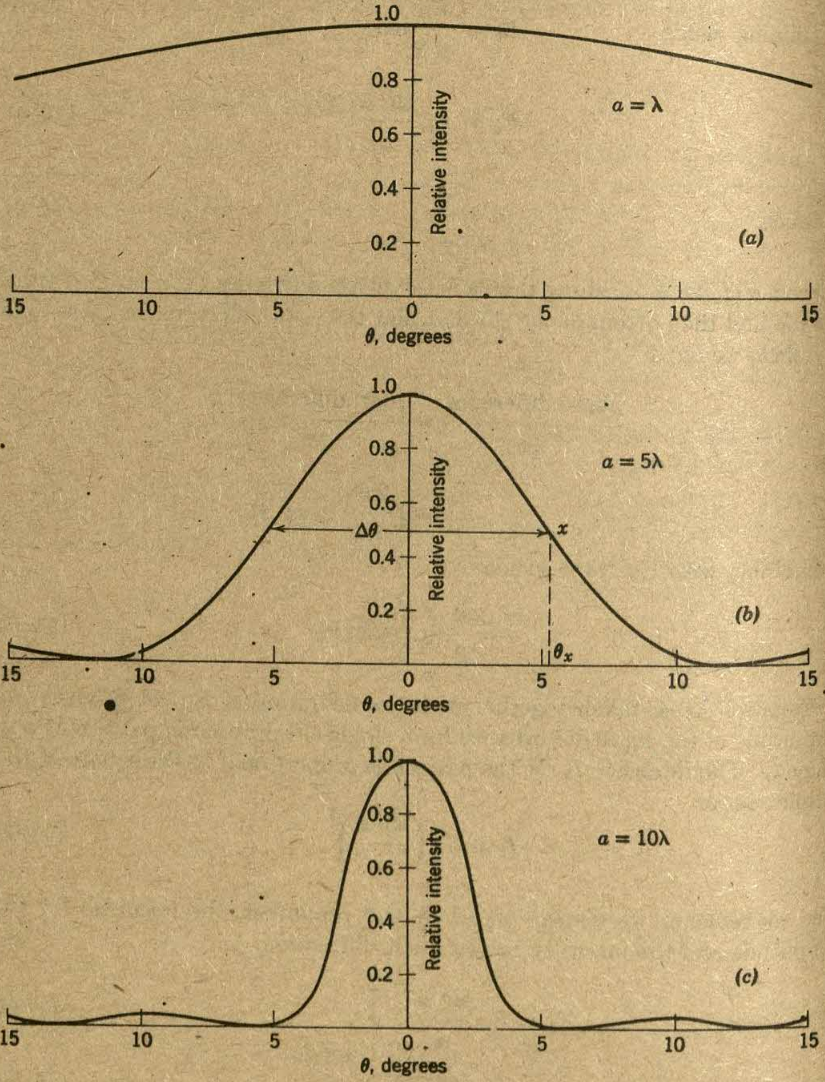


Fig. 44–9 The relative intensity in single-slit diffraction for three values of the ratio a/λ. The arrow in (*b*) shows the half-width $\Delta\theta$ of the central maximum.

Combining with Eq. 44–8*c* leads to

$$a \sin \theta = m\lambda \qquad m = 1, 2, 3, \ldots \text{ (minima)},$$

which is the result derived in the preceding section (Eq. 44–2). In that section, however, we derived *only* this result, obtaining no quantitative information about the intensity of the diffraction pattern at places in which it was not zero. Here (Eqs. 44–8) we have complete intensity information.

▶ **Example 3.** *Intensities of the secondary diffraction maxima.* Calculate, approximately, the relative intensities of the secondary maxima in the single-slit Fraunhofer diffraction pattern.

The secondary maxima lie approximately halfway between the minima and are found (compare Eq. 44-9) from

$$\alpha \cong (m + \tfrac{1}{2})\pi \qquad m = 1, 2, 3, \ldots .$$

Substituting into Eq. 44-8*b* yields

$$I_\theta = I_m \left[\frac{\sin (m + \frac{1}{2})\pi}{(m + \frac{1}{2})\pi}\right]^2,$$

which reduces to

$$\frac{I_\theta}{I_m} = \frac{1}{(m+\frac{1}{2})^2\pi^2}.$$

This yields, for $m = 1, 2, 3, \ldots, I_\theta/I_m = 0.045, 0.016, 0.0083$, etc. The successive maxima decrease rapidly in intensity.

Example 4. *Width of the central diffraction maximum.* Derive the *half-width* $\Delta\theta$ of the central maximum in a single-slit Fraunhofer diffraction (see Fig. 44-9*b*). The half-width is the angle between the two points in the pattern where the intensity is one-half that at the center of the pattern.

Point x in Fig. 44-9*b* is so chosen that $I_\theta = \frac{1}{2}I_m$, or, from Eq. 44-8*b*,

$$\frac{1}{2} = \left(\frac{\sin \alpha_x}{\alpha_x}\right)^2.$$

This equation cannot be solved analytically for α_x. It can be solved graphically, as accurately as one wishes, by plotting the quantity $(\sin \alpha_x/\alpha_x)^2$ as ordinate versus α_x as abscissa and noting the value of α_x at which the curve intersects the line "one-half" on the ordinate scale (see Problem 5). However, if only an approximate answer is desired, it is often quicker to use trial-and-error methods.

We know that α equals π at the first minimum; we guess that α_x is perhaps $\pi/2$ ($= 90° = 1.57$ radians). Trying this in Eq. 44-8*b* yields

$$\frac{I_\theta}{I_m} = \left[\frac{\sin (\pi/2)}{\pi/2}\right]^2 = 0.4.$$

This intensity ratio is *less* than 0.5, so that α_x must be *less* than 90°. After a few more trials we find easily enough that

$$\alpha_x = 1.40 \text{ radians} = 80°$$

does yield a ratio close to the correct value of 0.5.

We now use Eq. 44-8*c* to find the corresponding angle θ:

$$\alpha_x = \frac{\pi a}{\lambda} \sin \theta_x = 1.40,$$

or, noting that $a/\lambda = 5$ for Fig. 44-9*b*,

$$\sin \theta_x = \frac{1.40\lambda}{\pi a} = \frac{1.40}{5\pi} = 0.0892.$$

The half-width $\Delta\theta$ of the central maximum (see Fig. 44-9*b*) is given by

$$\Delta\theta = 2\theta_x = 2 \sin^{-1} 0.0892 = 2 \times 5.1° = 10.2°,$$

which is in agreement with the figure.

44-5 Diffraction at a Circular Aperture

Diffraction will occur when a wavefront is partially blocked off by an opaque object such as a metal disk or an opaque screen containing an aperture. Here we consider diffraction at a circular aperture of diameter d, the aperture constituting the boundary of a circular lens.

Our previous treatment of lenses was based on geometrical optics, diffraction being specifically assumed not to occur. A rigorous analysis would be based from the beginning on wave optics, since geometrical optics is always an approximation, although often a good one. Diffraction phenomena would emerge in a natural way from such a wave-optical analysis.

Figure 44-10 shows the image of a distant point source of light (a star) formed on a photographic film placed in the focal plane of a converging lens. It is not a point, as the (approximate) geometrical optics treatment suggests, but a circular disk surrounded by several progressively fainter secondary rings. Comparison with Fig. 42-3*c* leaves little doubt that we are dealing with a diffraction phenomenon in which, however, the aperture is a circle rather than a long narrow slit. The ratio d/λ, where d is the diameter of the lens (or of a circular aperture placed in front of the lens), determines the scale of the diffraction pattern, just as the ratio a/λ does for a slit.

Analysis shows that the first minimum for the diffraction pattern of a circular aperture of diameter d, assuming Fraunhofer conditions, is given by

$$\sin\theta = 1.22\frac{\lambda}{d}. \tag{44-10}$$

This is to be compared with Eq. 44-1, or

$$\sin\theta = \frac{\lambda}{a},$$

which locates the first minimum for a long narrow slit of width a. The factor 1.22 emerges from the mathematical analysis when we integrate over the elementary radiators into which the circular aperture may be divided.

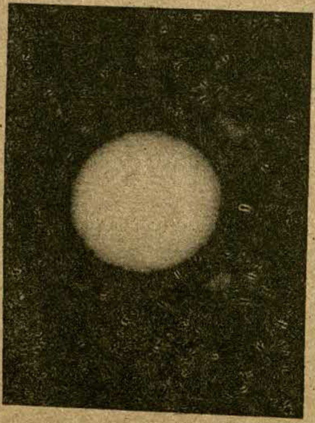


Fig. 44-10 The image of a star formed by a converging lens is a diffraction pattern. Note the central maximum, sometimes called the Airy disk (after Sir George Airy, who first solved the problem of diffraction at a circular aperture in 1835), and the circular secondary maximum. Other secondary maxima occur at larger radii but are too faint to be seen.

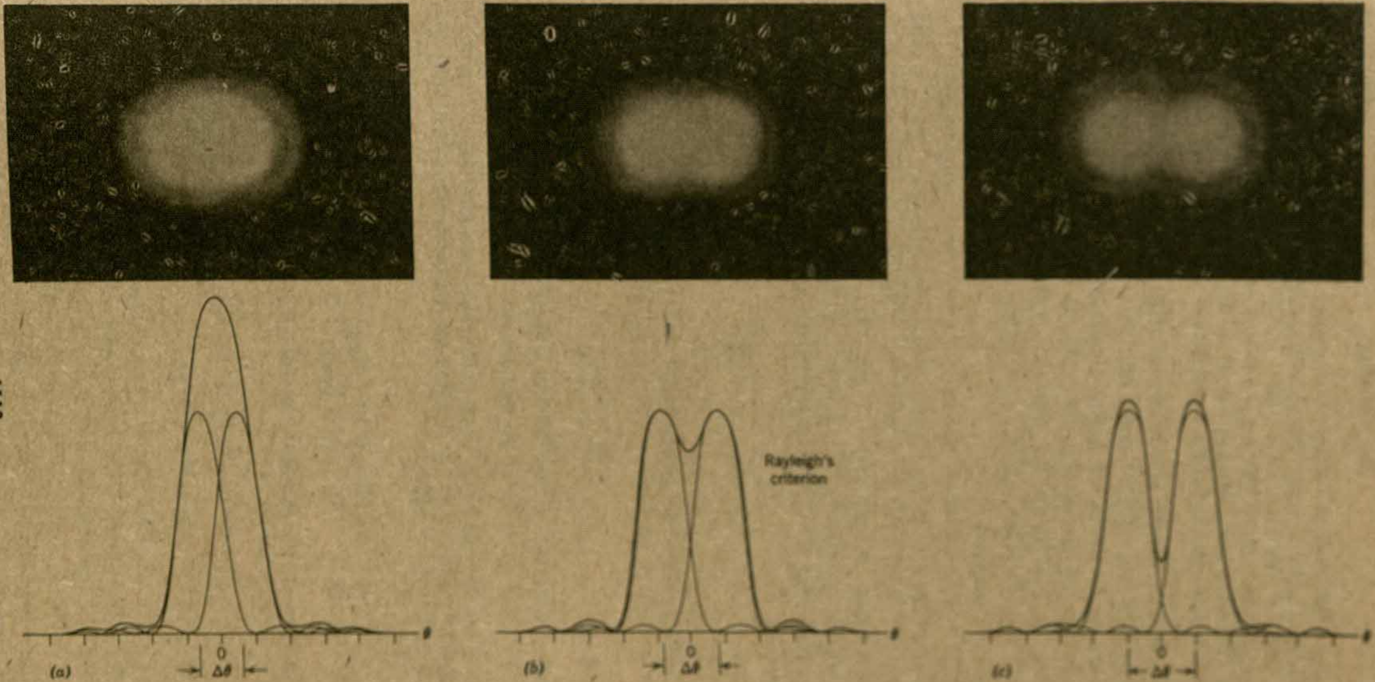


Fig. 44-11 The images of two distant point objects are formed by a converging lens whose diameter (= 10 cm) is 200,000 times the effective wavelength (= 5000 Å). Sketches of the images as they appear in the focal plane of the lens are shown with the corresponding intensity plots below them. (*a*) The angular separation of the objects (see vertical ticks) is so small that the images are not resolved. (*b*) The objects are farther apart and the images meet Rayleigh's criterion for resolution. (*c*) The objects are still farther apart and the images are well resolved.

In actual lenses the image of a distant point object will be somewhat larger than that shown in Fig. 44–10 and may not have radial symmetry. This is caused by the various lens "defects" mentioned on p. 1061. However, even if all of these defects could be eliminated by suitable shaping of the lens surfaces or by introducing correcting lenses, the diffraction pattern of Fig. 44–10 would remain. It is an inherent property of the lens aperture and of the wavelength of light used.

The fact that lens images are diffraction patterns is important when we wish to distinguish two distant point objects whose angular separation is small. Figure 44–11 shows the visual appearances and the corresponding intensity patterns for two distant point objects with small angular separations. In (*a*) the objects are not resolved; that is, they cannot be distinguished from a single point object. In (*b*) they are barely resolved and in (*c*) they are fully resolved.

In Fig. 44–11*b* the angular separation of the two point sources is such that the maximum of the diffraction pattern of one source falls on the first minimum of the diffraction pattern of the other. This is called *Rayleigh's criterion.* This criterion, though useful, is arbitrary; other criteria for deciding when two objects are resolved are sometimes used. From Eq. 44–10, two objects that are barely resolvable by Rayleigh's criterion must have an angular separation θ_R of

$$\theta_R = \sin^{-1} \frac{1.22\lambda}{d}.$$

Since the angles involved are rather small, we can replace $\sin \theta_R$ by θ_R, or

$$\theta_R = 1.22 \frac{\lambda}{d}. \tag{44–11}$$

If the angular separation θ between the objects is greater than θ_R, we can resolve the two objects; if it is less, we cannot.

▶ **Example 5.** A converging lens 3.0 cm in diameter has a focal length f of 20 cm. (*a*) What angular separation must two distant point objects have to satisfy Rayleigh's criterion? Assume that $\lambda = 5500$ A.

From Eq. 44–11,

$$\theta_R = 1.22 \frac{\lambda}{d} = \frac{(1.22)(5.5 \times 10^{-7} \text{ meter})}{3.0 \times 10^{-2} \text{ meter}} = 2.2 \times 10^{-5} \text{ radian}.$$

(*b*) How far apart are the centers of the diffraction patterns in the focal plane of the lens? The linear separation is

$$x = f\theta = (20 \text{ cm})(2.2 \times 10^{-5} \text{ radian}) = 44{,}000 \text{ A}.$$

This is 8.0 wavelengths of the light employed. ◀

When one wishes to use a lens to resolve objects of small angular separation, it is desirable to make the central disk of the diffraction pattern as small as possible. This can be done (see Eq. 44–11) by increasing the lens diameter or by using a shorter wavelength. One reason for constructing large telescopes is to produce *sharper* images so that celestial objects can be examined in finer detail. The images are also *brighter*, not only because the energy is concentrated into a smaller diffraction disk but because the larger lens collects more light. Thus fainter objects, for example, more distant stars, can be seen.

To reduce diffraction effects in *microscopes* we often use ultraviolet light, which, because of its shorter wavelength, permits finer detail to be examined than would be possible for the same microscope operated with visible light. We shall see in Chapter 48 that beams of electrons behave like waves under some circumstances. In the *electron microscope* such beams may have an effective wavelength of 0.04 A, of the order of 10^5 times shorter than visible light ($\lambda \cong 5000$ A). This permits the detailed examination of tiny objects like viruses. If a virus were examined with an optical microscope, its structure would be hopelessly concealed by diffraction.

44-6 Double Slit

In Young's double-slit experiment (Section 43-1) we assume that the slits are arbitrarily narrow (that is, $a \ll \lambda$), which means that the central part of the diffusing screen was uniformly illuminated by the diffracted waves from each slit. When such waves interfere, they produce fringes of uniform intensity, as in Fig. 43-9. This idealized situation cannot occur with actual slits because the condition $a \ll \lambda$ cannot usually be met. Waves from the two actual slits combining at different points of the screen will have intensities that are *not* uniform but are governed by the diffraction pattern of a single slit. The effect of relaxing the assumption that $a \ll \lambda$ in Young's experiment is to leave the fringes relatively unchanged in location but to alter their intensities.

The interference pattern for infinitesimally narrow slits is given by Eq. 43-11*b* and *c* or, with a small change in nomenclature,

$$I_{\theta,\text{int}} = I_{m,\text{int}} \cos^2 \beta, \tag{44-12}$$

where

$$\beta = \frac{\pi d}{\lambda} \sin \theta. \tag{44-13}$$

The intensity for the diffracted wave from either slit is given by Eqs. 44-8*b* and *c*, or, with a small change in nomenclature,

$$I_{\theta,\text{dif}} = I_{m,\text{dif}} \left(\frac{\sin \alpha}{\alpha}\right)^2, \tag{44-14}$$

where

$$\alpha = \frac{\pi a}{\lambda} \sin \theta. \tag{44-15}$$

The combined effect is found by regarding $I_{m,\text{int}}$ in Eq. 44-12 as a variable amplitude, given in fact by $I_{\theta,\text{dif}}$ of Eq. 44-14. This assumption, for the combined pattern, leads to

$$I_\theta = I_m (\cos \beta)^2 \left(\frac{\sin \alpha}{\alpha}\right)^2, \tag{44-16}$$

in which we have dropped all subscripts referring separately to interference and diffraction.

Let us express this result in words. At any point on the screen the available light intensity from each slit, considered separately, is given by the diffraction pattern of that slit (Eq. 44–14). The diffraction patterns for the two slits, again considered separately, coincide because parallel rays in

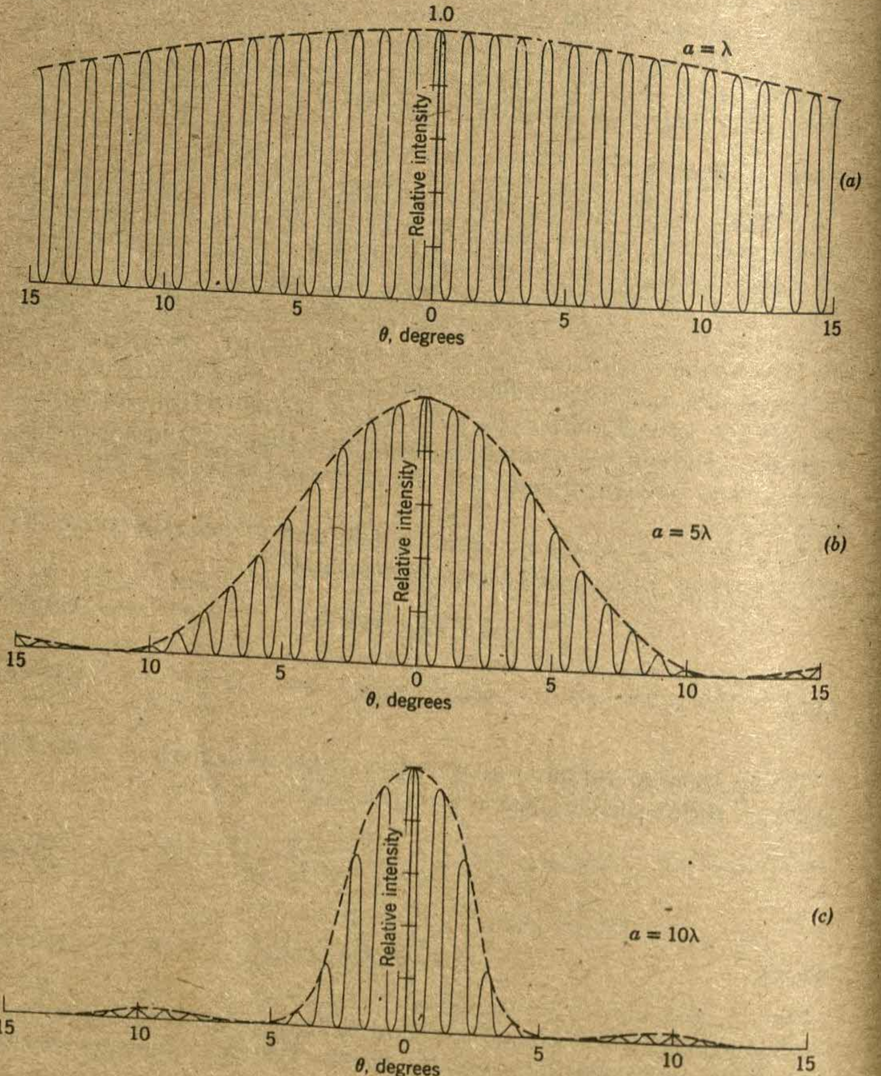


Fig. 44–12 Interference fringes for a double slit with slit separation $d = 50\lambda$. Three different slit widths, described by $a/\lambda = 1$, 5, and 10 are shown.

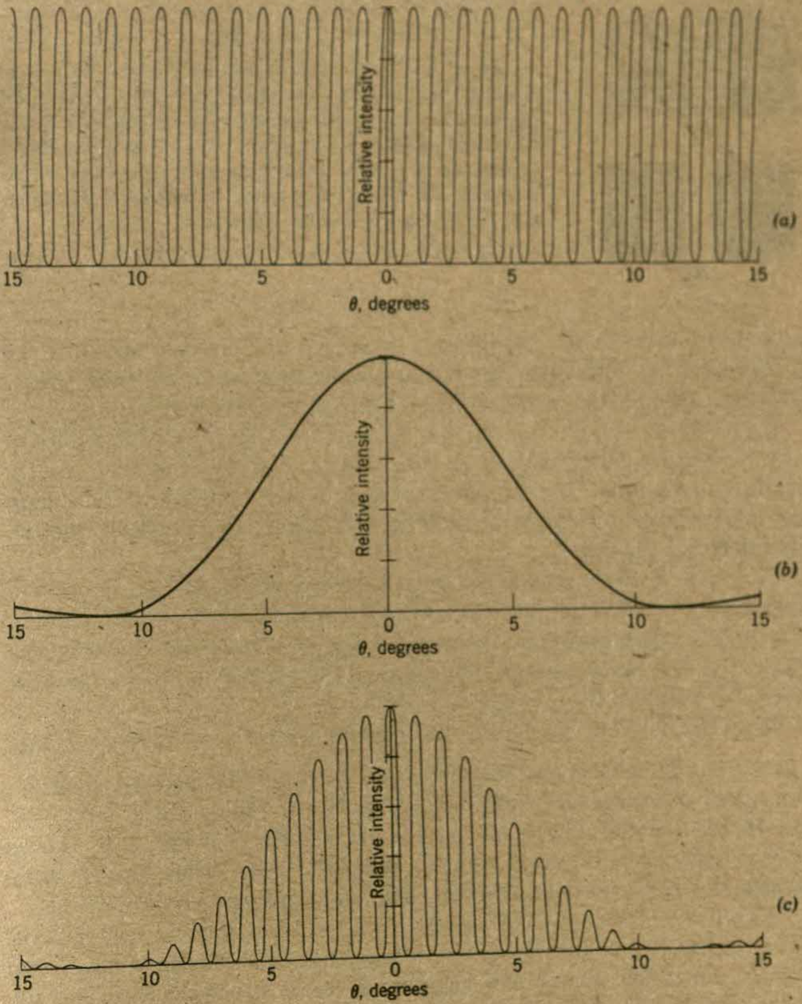
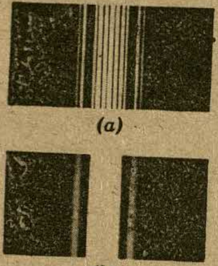


Fig. 44–13 (*a*) The "interference factor" and (*b*) the "diffraction factor" in Eq. 44–16 and (*c*) their product; compare Fig. 44–12*b*.

Fraunhofer diffraction are focused at the same spot (see Fig. 44–5). Because the two diffracted waves are coherent, they will interfere.

The effect of interference is to redistribute the available energy over the screen, producing a set of fringes. In Section 43–1, where we assumed $a \ll \lambda$, the available energy was virtually the same at all points on the screen so that the interference fringes had virtually the same intensities (see Fig. 43–9). If we relax the assumption $a \ll \lambda$, the available energy is *not* uniform over the screen but is given by the diffraction pattern of a slit of width a. In this


(a)

(b)

Fig. 44–14 (*a*) Interference fringes for a double-slit system in which the slit width is *not* negligible in comparison to the wavelength. The fringes are modulated in intensity by the diffraction pattern of a single slit. (*b*) If one of the slits is covered up, the interference fringes disappear and we see the single slit diffraction pattern. (Courtesy G. H. Carragan, Rensselaer Polytechnic Institute.)

case the interference fringes will have intensities that are determined by the intensity of the diffraction pattern at the location of a particular fringe. Equation 44–16 is the mathematical expression of this argument.

Figure 44–12 is a plot of Eq. 44–16 for $d = 50\lambda$ and for three values of a/λ. It shows clearly that for narrow slits ($a = \lambda$) the fringes are nearly uniform in intensity. As the slits are widened, the intensities of the fringes are markedly modulated by the "diffraction factor" in Eq. 44–16, that is, by the factor $(\sin \alpha/\alpha)^2$.

Equation 44–16 shows that the fringe envelopes of Fig. 44–12 are precisely the single-slit diffraction patterns of Fig. 44–9. This is especially clear in Fig. 44–13, which shows, for the curve of Fig. 44–12*b*, (*a*) the "interference factor" in Eq. 44–16 (that is, the factor $\cos^2 \beta$), (*b*) the "diffraction factor" $(\sin \alpha/\alpha)^2$, and (*c*) their product.

If we put $a = 0$ in Eq. 44–16, then (see Eq. 44–15) $\alpha = 0$ and $\sin \alpha/\alpha \underset{\alpha\to 0}{\cong} \alpha/\alpha = 1$. Thus this equation reduces, as it must, to the intensity equation for a pair of vanishingly narrow slits (Eq. 44–12). If we put $d = 0$ in Eq. 44–16, the two slits coalesce into a single slit of width a, as Fig. 44–15 shows; $d = 0$ implies $\beta = 0$ (see Eq. 44–13) and $\cos^2 \beta = 1$. Thus Eq. 44–16 reduces, as it must, to the diffraction equation for a single slit (Eq. 44–14).

Figure 44–14 shows some actual double-slit interference photographs. The uniformly spaced interference fringes and their intensity modulation by the diffraction pattern of a single slit is clear. If one slit is covered up, as in Fig. 44–14*b*, the interference fringes disappear and we see the diffraction pattern of a single slit.

▶ **Example 6.** Starting from the curve of Fig. 44–12*b*, what is the effect of (*a*) increasing the slit width, (*b*) increasing the slit separation, and (*c*) increasing the wavelength?

(*a*) If we increase the slit width a, the envelope of the fringe pattern changes so that its central peak is sharper (compare Fig. 44–12*c*). The fringe spacing, which depends on d/λ, does not change.

(*b*) If we increase d, the fringes become closer together, the envelope of the pattern remaining unchanged.

(*c*) If we increase λ, the envelope becomes broader and the fringes move further apart. Increasing λ is equivalent to decreasing both of the ratios a/λ and d/λ. The general relationship of the envelope to the fringes, which depends only on d/a, does not change with wavelength.

Example 7. In double-slit Fraunhofer diffraction what is the fringe spacing on a screen 50 cm away from the slits if they are illuminated with blue light ($\lambda = 4800$ A), if $d = 0.10$ mm, and if the slit width $a = 0.02$ mm? What is the linear distance from the central maximum to the first minimum of the fringe envelope?

The intensity pattern is given by Eq. 44-16, the fringe spacing being determined by the interference factor $\cos^2 \beta$. From Example 2, Chapter 43, we have

$$\Delta y = \frac{\lambda D}{d},$$

where D is the distance of the screen from the slits. Substituting yields

$$\Delta y = \frac{(480 \times 10^{-9} \text{ meter})(50 \times 10^{-2} \text{ meter})}{0.10 \times 10^{-3} \text{ meter}} = 2.4 \times 10^{-3} \text{ meter} = 2.4 \text{ mm}.$$

The distance to the first minimum of the envelope is determined by the diffraction factor $(\sin \alpha/\alpha)^2$ in Eq. 44-16. The first minimum in this factor occurs for $\alpha = \pi$. From Eq. 44-15,

$$\sin \theta = \frac{\alpha\lambda}{\pi a} \doteq \frac{\lambda}{a} = \frac{480 \times 10^{-9} \text{ meter}}{0.02 \times 10^{-3} \text{ meter}} = 0.024.$$

This is so small that we can assume that $\theta \cong \sin \theta \cong \tan \theta$, or

$$y = D \tan \theta \cong D \sin \theta = (50 \text{ cm})(0.024) = 1.2 \text{ cm}.$$

There are about ten fringes in the central peak of the fringe envelope.

Example 8. What requirements must be met for the central maximum of the envelope of the double-slit Fraunhofer pattern to contain exactly eleven fringes?

The required condition will be met if the sixth minimum of the interference factor ($\cos^2 \beta$) in Eq. 44-16 coincides with the first minimum of the diffraction factor $(\sin \alpha/\alpha)^2$.

The sixth minimum of the interference factor occurs when

$$\beta = \tfrac{11}{2}\pi$$

in Eq. 44-12.

The first minimum in the diffraction term occurs for

$$\alpha = \pi.$$

Dividing (see Eqs. 44-13 and 44-15) yields

$$\frac{\beta}{\alpha} = \frac{d}{a} = \frac{11}{2}.$$

This condition depends only on the slit geometry and not on the wavelength. For long waves the pattern will be broader than for short waves, but there will always be eleven fringes in the central peak of the envelope. ◀

The double-slit problem as illustrated in Fig. 44-12 combines interference and diffraction in an intimate way. At root both are superposition effects and depend on adding wave disturbances at a given point, taking phase differences properly into account. If the waves to be combined originate from a *finite* (and usually small) number of elementary coherent radiators, as in Young's double-slit experiment, we call the effect *interference*. If the waves to be combined originate by subdividing a wave into *infinitesimal* coherent radiators, as in our treatment of a single slit (Fig. 44–6), we call the effect *diffraction*. This distinction between interference and diffraction is convenient and useful. However, it should not cause us to lose sight of the fact that both are superposition effects and that often both are present simultaneously, as in Young's experiment.

QUESTIONS

1. Why is the diffraction of sound waves more evident in daily experience than that of light waves?

2. Why do radio waves diffract around buildings, although light waves do not?

3. A loud-speaker horn has a rectangular aperture 4 ft high and 1 ft wide. Will the pattern of sound intensity be broader in the horizontal plane or in the vertical?

4. A radar antenna is designed to give accurate measurements of the height of an aircraft but only reasonably good measurements of its direction in a horizontal plane. Must the height-to-width ratio of the radar reflector be less than, equal to, or greater than unity?

5. A person holds a single narrow vertical slit in front of the pupil of his eye and looks at a distant light source in the form of a long heated filament. Is the diffraction pattern that he sees a Fresnel or a Fraunhofer pattern?

6. In a single-slit Fraunhofer diffraction, what is the effect of increasing (*a*) the wavelength and (*b*) the slit width?

7. Sunlight falls on a single slit of width 10^4 A. Describe qualitatively what the resulting diffraction pattern looks like.

8. In Fig. 44–5 rays r_1 and r_3 are in phase; so are r_2 and r_4. Why isn't there a *maximum* intensity at P_2 rather than a minimum?

9. Describe what happens to a Fraunhofer single-slit diffraction pattern if the whole apparatus is immersed in water.

10. Distinguish clearly between θ, α, and ϕ in Eq. 44–8*c*.

11. Do diffraction effects occur for virtual images as well as for real images? Explain.

12. Do diffraction effects occur for images formed by (*a*) plane mirrors and (*b*) spherical mirrors? Explain.

13. If we were to redo our analysis of the properties of lenses in Section 42–5 by the methods of geometrical optics but *without* restricting our considerations to paraxial rays and to "thin" lenses, would diffraction phenomena, such as that of Fig. 44–10, emerge from the analysis? Discuss.

14. Distinguish carefully between interference and diffraction in Young's double-slit experiment.

15. In what way are interference and diffraction similar? In what way are they different?

16. In double-slit interference patterns such as that of Fig. 44–14*a* we said that the interference fringes were modulated in intensity by the diffraction pattern of a single slit. Could we reverse this statement and say that the diffraction pattern of a single slit is intensity-modulated by the interference fringes? Discuss.

PROBLEMS

1. In a single-slit diffraction pattern the distance between the first minimum on the right and the first minimum on the left is 5.2 mm. The screen on which the pattern is displayed is 80 cm from the slit and the wavelength is 5460 A. Calculate the slit width.

2. A plane wave ($\lambda = 5900$ A) falls on a slit with $a = 0.40$ mm. A converging lens ($f = +70$ cm) is placed behind the slit and focuses the light on a screen. What is the linear distance on the screen from the center of the pattern to (*a*) the first minimum and (*b*) the second minimum?

3. A single slit is illuminated by light whose wavelengths are λ_a and λ_b, so chosen that the first diffraction minimum of λ_a coincides with the second minimum of λ_b. (*a*) What relationship exists between the two wavelengths? (*b*) Do any other minima in the two patterns coincide?

4. (*a*) Show that the values of α at which intensity maxima for single-slit diffraction occur can be found exactly by differentiating Eq. 44–8*b* with respect to α and equating to zero, obtaining the condition

$$\tan \alpha = \alpha.$$

(*b*) Find the values of α satisfying this relation by plotting graphically the curve $y = \tan \alpha$ and the straight line $y = \alpha$ and finding their intersections. (*c*) Find the (nonintegral) values of m corresponding to successive maxima in the single-slit pattern. Note that the secondary maxima do not lie exactly halfway between minima.

5. In Example 4 solve the transcendental equation

$$\frac{1}{2} = \left(\frac{\sin \alpha_x}{\alpha_x}\right)^2$$

graphically for α_x, to an accuracy of three significant figures.

6. (*a*) In Fig. 44–7*d*, why is E_θ, which represents the first maximum beyond the central maximum, not vertical? (*b*) Calculate the angle it makes with the vertical, assuming the slit to be divided into infinitesimal strips of width dx.

7. What is the half-width of a diffracted beam for a slit whose width is (*a*) 1, (*b*) 5, and (*c*) 10 wavelengths?

8. (*a*) A circular diaphragm 0.60 meter in diameter oscillates at a frequency of 25,000 cycles/sec in an underwater source of sound for submarine detection. Far from the source the sound intensity is distributed as a Fraunhofer diffraction pattern for a circular hole whose diameter equals that of the diaphragm. Take the speed of sound in water to be 1450 meters/sec and find the angle between the normal to the diaphragm and the direction of the first minimum. (*b*) Repeat for a source having an (audible) frequency of 1000 cycles/sec.

9. The two headlights of an approaching automobile are 4 ft apart. At what maximum distance will the eye resolve them? Assume a pupil diameter of 5.0 mm and $\lambda = 5500$ A. Assume also that this distance is determined only by diffraction effects at the circular pupil aperture.

10. The wall of a large room is covered with acoustic tile in which small holes are drilled 5.0 mm from center to center. How far can a person be from such a tile and still distinguish the individual holes, assuming ideal conditions? Assume the diameter of the pupil to be 4.0 mm and λ to be 5500 A.

11. (*a*) How small is the angular separation of two stars if their images are barely resolved by the Thaw refracting telescope at the Allegheny Observatory in Pittsburgh? The lens diameter is 30 in. and its focal length is 46 ft. Assume $\lambda = 5000$ A. (*b*) Find the distance between these barely resolved stars if each of them is 10 light years distant from the earth. (*c*) For the image of a single star in this telescope, find the diameter of the first dark ring in the diffraction pattern, as measured on a photographic plate placed at the

focal plane. Assume that the star image structure is associated entirely with diffraction at the lens aperture and not with (small) lens "errors."

12. Find the separation of two points on the moon's surface that can just be resolved by the 200-in. telescope at Mount Palomar, assuming that this distance is determined by diffraction effects. The distance from the earth to the moon is 240,000 miles.

13. Construct qualitative vector diagrams like those of Fig. 44–7 for the double-slit interference pattern. For simplicity, consider $d = 2a$ (see Fig. 44–15). Can you interpret the main features of the intensity pattern this way?

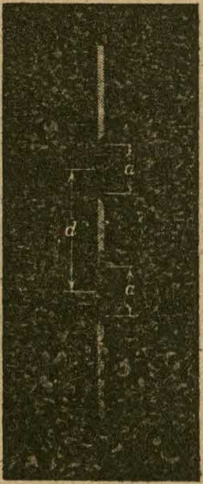


Fig. 44–15

14. Suppose that, as in Example 8, the envelope of the central peak contains eleven fringes. How many fringes lie between the first and second minima of the envelope?

15. For $d = 2a$ in Fig. 44–15, how many interference fringes lie in the central diffraction envelope?

16. If we put $d = a$ in Fig. 44–15, the two slits coalesce into a single slit of width $2a$. Show that Eq. 44–16 reduces to the diffraction pattern for such a slit.

17. (*a*) Design a double-slit system in which the fourth fringe, not counting the central maximum, is missing. (*b*) What other fringes, if any, are also missing?

Gratings and Spectra

CHAPTER 45

45–1 Introduction

In connection with Young's experiment (Sections 43–1 and 43–3) we discussed the interference of two coherent waves formed by diffraction at two elementary radiators (pinholes or slits). In our first treatment we assumed that the slit width was much less than the wavelength, so that light diffracted from each slit illuminated the observation screen essentially uniformly. Later, in Section 44–6, we took the slit width into account and showed that the intensity pattern of the interference fringes is modulated by a "diffraction factor" $(\sin \alpha/\alpha)^2$ (see Eq. 44–16).

Here we extend our treatment to cases in which the number N of radiators or diffracting centers is larger—and usually much larger—than two. We consider two situations:

1. An array of N parallel equidistant slits, called a *diffraction grating.*
2. A three-dimensional array of periodically arranged radiators—the atoms in a crystalline solid such as NaCl. In this case the average spacing between the elementary radiators is so small that interference effects must be sought at wavelengths much smaller than those of visible light. We speak of *X-ray diffraction.*

In each case we distinguish carefully between the diffracting properties of a single radiator (slit or atom) and the interference of the waves diffracted, coherently, from the assembly of radiators.

45–2 Multiple Slits

A logical extension of Young's double-slit interference experiment is to increase the number of slits from two to a larger number N. An arrangement like that of Fig. 45–1, usually involving many more slits, is called a *diffrac-*

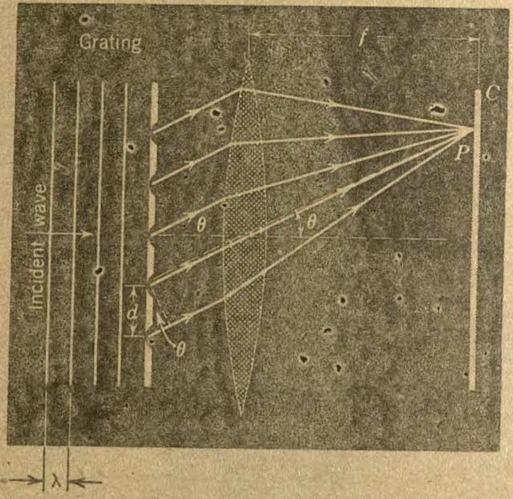


Fig. 45-1 An idealized diffraction grating containing five slits. The slit width a is shown for convenience to be considerably smaller than λ, although this condition is not realized in practice. The figure is distorted in that f is much greater than d in practice.

tion grating. As for a double slit, the intensity pattern that results when monochromatic light of wavelength λ falls on a grating consists of a series of interference fringes. The *angular separations* of these fringes are determined by the ratio λ/d, where d is the spacing between the centers of adjacent slits. The relative *intensities* of these fringes are determined by the diffraction pattern of a single grating slit, which depends on the ratio λ/a, where a is the slit width.

Figure 45-2, which compares the intensity patterns for $N = 2$ and $N = 5$, shows clearly that the "interference" fringes are modulated in intensity by a "diffraction" envelope, as in Fig. 44-14. Figure 45-3 presents a theoretical calculation of the intensity patterns for a few fringes near the centers of the patterns of Fig. 45-2. These two figures show that increasing N (a) does not change the spacing between the (principal) interference fringe maxima,

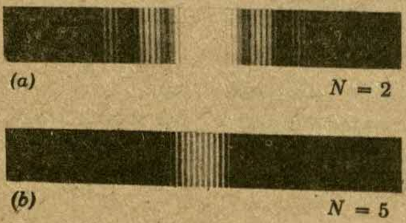


Fig. 45-2 Intensity patterns for "gratings" with (a) $N = 2$ and (b) $N = 5$ for the same value of d and λ. Note how the intensities of the fringes are modulated by a diffraction envelope as in Fig. 44-14; thus the assumption $a \ll \lambda$ is not realized in these actual "gratings." For $N = 5$ three very faint secondary maxima, not visible in this photograph, appear between each pair of adjacent primary maxima.

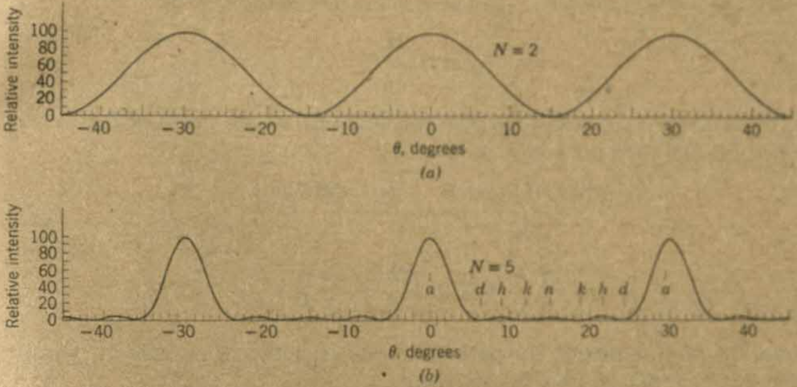


Fig. 45-3 Calculated intensity patterns for (*a*) a two-slit and (*b*) a five-slit grating for the same value of d and λ. This figure shows the sharpening of the principal maxima and the appearance of faint secondary maxima for $N > 2$. The letters on the five-slit pattern refer to Fig. 45-5. The figure assumes slits with $a \ll \lambda$ so that the principal maxima are of uniform intensity.

provided d and λ remain unchanged, (*b*) sharpens the (principal) maxima, and (*c*) introduces small secondary maxima between the principal maxima. Three such secondaries are present (but not readily visible) between each pair of adjacent principal maxima in Fig. 45-2*b*.

A principal maximum in Fig. 45-1 will occur when the path difference between rays from adjacent slits ($= d \sin \theta$) is given by

$$d \sin \theta = m\lambda \qquad m = 0, 1, 2, \ldots \qquad \text{(principal maxima)}, \qquad (45\text{-}1)$$

where m is called the *order number*. This equation is identical with Eq. 43-1, which locates the intensity maxima for a double slit. The *locations* of the (principal) maxima are thus determined only by the ratio λ/d and are independent of N. As for the double slit, the ratio a/λ determines the relative *intensities* of the principal maxima but does not alter their locations appreciably.

The sharpening of the principal maxima as N is increased can be understood by a graphical argument, using phasors. Figures 45-4*a* and *b* show conditions at any of the principal maxima for a two-slit and a nine-slit grating. The small arrows represent the amplitudes of the wave disturbances arriving at the screen at the position of each principal maximum. For simplicity we consider the central principal maximum only, for which $m = 0$, and thus $\theta = 0$, in Eq. 45-1.

Consider the angle $\Delta\theta_0$ corresponding to the position of zero intensity that lies on either side of the central principal maximum. Figures 45-4*c* and *d* show the phasors at this point. The phase difference between waves from adjacent slits, which is zero at the central principal maximum, must increase by an amount $\Delta\phi$ chosen so that the array of phasors just closes on itself, yielding zero resultant intensity. For $N = 2$, $\Delta\phi = 2\pi/2$ ($= 180°$); for $N = 9$, $\Delta\phi = 2\pi/9$ ($= 40°$). In the general case it is given by

$$\Delta\phi = \frac{2\pi}{N}.$$

This increase in phase difference for adjacent waves corresponds to an increase in the path difference Δl given by

$$\frac{\text{phase difference}}{2\pi} = \frac{\text{path difference}}{\lambda},$$

or

$$\Delta l = \left(\frac{\lambda}{2\pi}\right)\Delta\phi = \left(\frac{\lambda}{2\pi}\right)\left(\frac{2\pi}{N}\right) = \frac{\lambda}{N}.$$

From Fig. 45–1, however, the path difference Δl at the first minimum is also given by $d \sin \Delta\theta_0$, so that we can write

$$d \sin \Delta\theta_0 = \frac{\lambda}{N},$$

or

$$\sin \Delta\theta_0 = \frac{\lambda}{Nd}.$$

Since $N \gg 1$ for actual gratings, $\sin \Delta\theta_0$ will ordinarily be quite small (that is, the lines will be sharp), and we may replace it by $\Delta\theta_0$ to good approximation, or

$$\Delta\theta_0 = \frac{\lambda}{Nd} \quad \text{(central principal maximum).} \tag{45–2}$$

This equation shows specifically that if we increase N for a given λ and d,

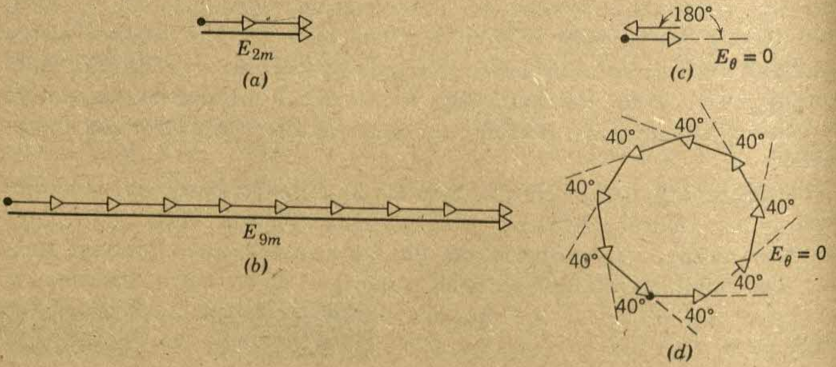


Fig. 45–4 Drawings (*a*) and (*b*) show conditions at the central principal maximum for a two-slit and a nine-slit grating, respectively. Drawings (*c*) and (*d*) show conditions at the minimum of zero intensity that lies on either side of this central principal maximum. In going from (*a*) to (*c*) the phase shift between waves from adjacent slits changes by 180° ($\Delta\phi = 2\pi/2$); in going from (*b*) to (*d*) it changes by 40° ($\Delta\phi = 2\pi/9$).

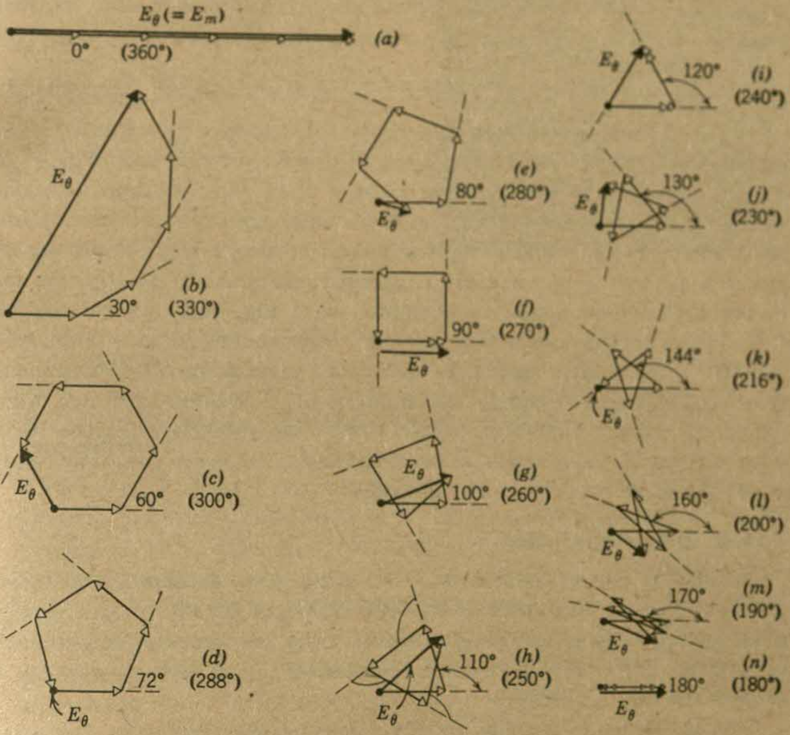


Fig. 45–5 The figures taken in sequence from (*a*) to (*n*) and then from (*n*) to (*a*) show conditions as the intensity pattern of a five-slit grating is traversed from the central principal maximum to an adjacent principal maximum. Phase differences between waves from adjacent slits are shown directly or, when going from (*n*) to (*a*), in parentheses. Principal maxima occur at (*a*), secondary maxima at, or near, (*h*) and (*n*), and points of zero intensity at (*d*) and (*k*). Compare Fig. 45–3*b*.

$\Delta\theta_0$ will decrease, which means that the central principal maximum becomes sharper.

We state without proof,* and for later use, that for principal maxima other than the central one (that is, for $m \neq 0$) the angular distance between the position θ_m of the principal maximum of order m and the minimum that lies on either side is given by

$$\Delta\theta_m = \frac{\lambda}{Nd \cos\theta_m} \qquad \text{(any principal maximum).} \tag{45-3}$$

For the central principal maximum we have $m = 0$, $\theta_m = 0$, and $\Delta\theta_m = \Delta\theta_0$, so that Eq. 45-3 reduces, as it must, to Eq. 45-2.

The origin of the secondary maxima that appear for $N > 2$ can also be understood using the phasor method. Figure 45-5*a* shows conditions at the central principal maximum for a five-slit grating. The vectors are in phase. As we depart from the central maximum, θ in Fig. 45–1 increases from zero and the phase difference between adjacent vectors increases from zero to $\Delta\phi = \frac{2\pi}{\lambda} d \sin \theta$. Successive figures show how the resultant wave amplitude E_θ varies with $\Delta\phi$. The student should verify by graphical construction that a given figure represents conditions for both $\Delta\phi$ and $2\pi - \Delta\phi$. Thus we start at $\Delta\phi = 0$, proceed to $\Delta\phi = 180°$, and then trace backward through the sequence, following the phase differences shown in parentheses, until we reach $\Delta\phi = 360°$. This sequence corresponds to traversing the intensity pattern from the central principal maximum to an adjacent one. Figure 45–5, which should be compared with Fig. 45–3*b*, shows that for $N = 5$ there are three secondary maxima, corresponding to $\Delta\phi = 110°, 180°$, and 250°. The student should make a similar analysis for $N = 3$ and should show that only one secondary maximum occurs. In actual gratings, which commonly contain 10,000 to 50,000 "slits," the secondary maxima lie so close to the principal maxima or are so reduced in intensity that they cannot be distinguished from them experimentally.

---

* See Problem 15.

45–3 Diffraction Gratings

The *grating spacing d* for a typical grating that contains 12,000 "slits" distributed over a 1-in. width is 2.54 cm/12,000, or 21,000 A. Gratings are often used to measure wavelengths and to study the structure and intensity of spectrum lines. Few devices have contributed more to our knowledge of modern physics.

Gratings are made by ruling equally spaced parallel grooves on a glass or a metal * plate, using a diamond cutting point whose motion is automatically controlled by an elaborate ruling engine. Once such a master grating has been prepared, replicas can be formed by pouring a collodion solution on the grating, allowing it to harden, and stripping it off. The stripped collodion, fastened to a flat piece of glass or other backing, forms a good grating.

Figure 45–6 shows a cross section of a common type of grating ruled on glass. In the rudimentary grating of Fig. 45–1 open slits were separated by opaque strips; the ***amplitude*** of the wave disturbance varied in a periodic way as the grating was crossed, dropping to zero on the opaque strips. The grating of Fig. 45–6 is transparent everywhere, so that there is little periodic change in amplitude as the grating is crossed. The effect of the rulings is to change the *optical thickness* of the grating in a periodic way, rays traversing the grating between the rulings (*b* in Fig. 45–6) containing more wavelengths than rays traversing the grating in the center of the rulings (*a* in Fig. 45–6). This results in a periodic change of *phase* as one crosses the grating at right angles to the rulings. Reflection gratings also depend for their operation on a periodic change in phase of the reflected wave as one crosses the grating, the change in amplitude under these conditions being negligible. The principal maxima for *phase gratings*, assuming that the incident light falls on the grating at right angles, can be given by the same formula derived earlier for idealized amplitude or slit gratings,

---

* Gratings ruled on metal are called *reflection gratings* because the interference effects are viewed in reflected rather than in transmitted light. Many research gratings are the reflection type; commonly they are ruled on the surface of a concave mirror, which eliminates the need for lenses.

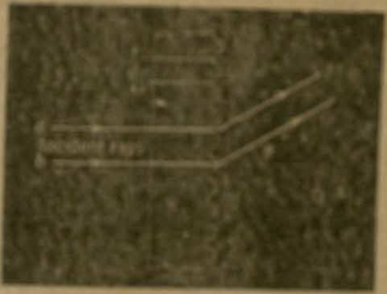


Fig. 45-6 An enlarged cross section of a diffraction grating ruled on glass. Such gratings, in which the phase of the emerging wave changes as one crosses the grating, are called *phase gratings*.

namely

$$d \sin \theta = m\lambda \qquad m = 0, 1, 2, \ldots$$

where d is the distance between the rulings and the integer m is called the *order of* the particular principal maximum. Essentially all gratings used in the visible spectrum, whether of the transmission type, as in Fig. 45-6, or the reflection type, are phase gratings.

Figure 45-7 shows a simple grating spectroscope, used for viewing the spectrum of a light source, assumed to emit a number of discrete wavelengths, or *spectrum lines*. The light from source S is focused by lens L_1 on a slit S_1 placed in the focal plane of lens L_2. The parallel light emerging from collimator C falls on grating G. Parallel rays associated with a particular interference maximum occurring at angle θ fall on lens L_2, being brought to a focus in plane F-F'. The image formed in this plane is examined, using a magnifying lens arrangement E, called an eyepiece. A symmetrical inter-

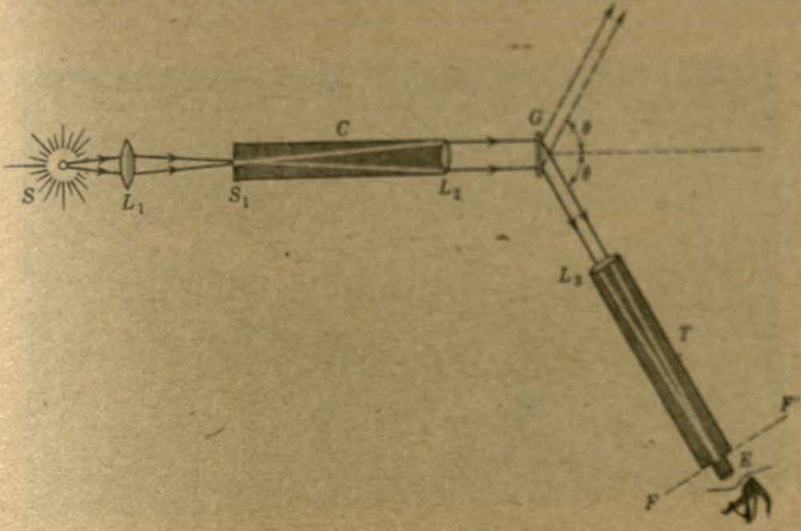


Fig. 45-7 A simple type of grating spectroscope used to analyze the wavelengths of the light emitted by source S.

ference pattern is formed on the other side of the central position, as shown by the dotted lines. The entire spectrum can be viewed by rotating telescope T through various angles. Instruments used for scientific research or in industry are more complex than the simple arrangement of Fig. 45–7. They invariably employ photographic or photoelectric recording and are called *spectrographs.* Figure 47–12 shows a small portion of the spectrum of iron, produced by examining the light produced in an arc struck between iron electrodes, using a research type spectrograph with photographic recording. Each line in the figure represents a different wavelength that is emitted from the source.

Grating instruments can be used to make absolute measurements of wavelength, since the grating spacing d in Eq. 45–1 can be measured accurately with a traveling microscope. Several spectra are normally produced in such instruments, corresponding to $m = \pm 1, \pm 2$, etc., in Eq. 45–1 (see Fig. 45–8). This may cause some confusion if the spectra overlap. Further, this multiplicity of spectra reduces the recorded intensity of any given spectrum line because the available energy is divided among a number of spectra.

This disadvantage of the grating instrument can be overcome by shaping the profile of the grating grooves so that a large fraction of the light is thrown into a particular order on a particular side (for a given wavelength). This technique, called *blazing,* so alters the diffracting properties of the individual grooves (by controlling their profiles) that the light of wavelength λ diffracted by a single groove has a sharp peak of maximum intensity at a selected angle θ ($\neq 0$).

Light can also be analyzed into its component wavelengths if the grating in Fig. 45–7 is replaced by a prism. In a *prism spectrograph* each wavelength

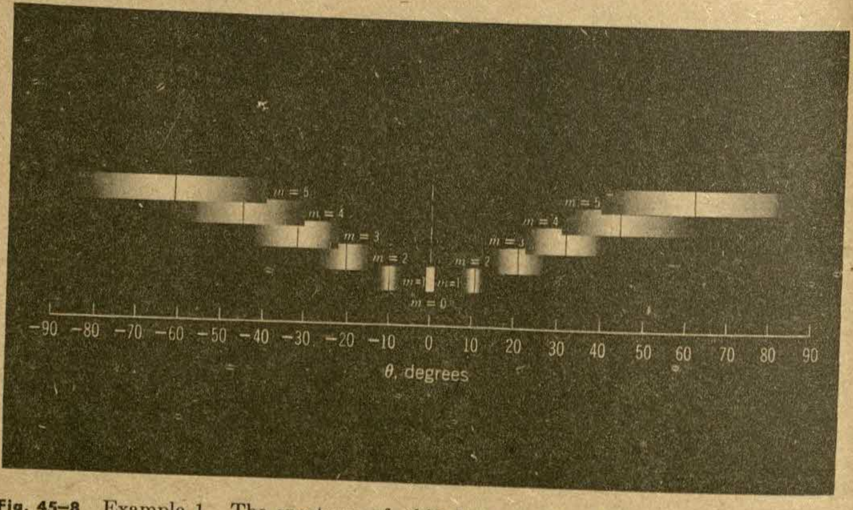


Fig. 45–8 Example 1. The spectrum of white light as viewed in a grating instrument like that of Fig. 45–7. The different orders, identified by the order number m, are shown separated vertically for clarity. As actually viewed, they would not be so displaced. The central line in each order corresponds to λ = 5500 A.

in the incident beam is deflected through a definite angle θ, determined by the index of refraction of the prism material for that wavelength. Curves such as Fig. 41-2, which gives the index of refraction of fused quartz as a function of wavelength, show that the shorter the wavelength, the larger the angle of deflection θ. Such curves vary from substance to substance and must be found by measurement. Prism instruments are not adequate for accurate *absolute* measurements of wavelength because the index of refraction of the prism material at the wavelength in question is usually not known precisely enough. Both prism and grating instruments make accurate *comparisons* of wavelength, using a suitable comparison spectrum such as that shown in Fig. 47-12, in which careful absolute determinations have been made of the wavelengths of the spectrum lines. The prism instrument has an advantage over a grating instrument (unblazed) in that its light energy is concentrated into a single spectrum so that brighter lines may be produced.

▶ **Example 1.** A grating with 8000 rulings/in. is illuminated with white light at perpendicular incidence. Describe the diffraction pattern. Assume that the wavelength of the light extends from 4000 to 7000 A.

The grating spacing d is 2.54 cm/8000, or 31,700 A. The central or zero-order maximum corresponds to $m = 0$ in Eq. 45-1. All wavelengths present in the incident light are superimposed at $\theta = 0$, as Fig. 45-8 shows.

The first-order diffraction pattern corresponds to $m = 1$ in Eq. 45-1. The 4000-A line occurs at an angle given by

$$\theta = \sin^{-1}\frac{m\lambda}{d} = \sin^{-1}\frac{(1)(4000\text{ A})}{31{,}700\text{ A}} = \sin^{-1} 0.126 = 7.3°.$$

In the same way the angle for the 7000-A line is found to be 12.8°, and the entire pattern of Fig. 45-8 can be calculated. Note that the *first-order spectrum* ($m = 1$) is isolated but that the second-, third-, and fourth-order spectra overlap.

Example 2. A diffraction grating has 10^4 rulings uniformly spaced over 1 in. It is illuminated at normal incidence by yellow light from a sodium vapor lamp. This light contains two closely spaced lines (the well-known *sodium doublet*) of wavelengths 5890.0 and 5895.9 A. (*a*) At what angle will the first-order maximum occur for the first of these wavelengths?

The grating spacing d is 10^{-4} in., or 25,400 A. The first-order maximum corresponds to $m = 1$ in Eq. 45-1. We thus have

$$\theta = \sin^{-1}\frac{m\lambda}{d} = \sin^{-1}\frac{(1)(5890\text{ A})}{25{,}400\text{ A}} = \sin^{-1} 0.232 = 13.3°.$$

(*b*) What is the angular separation between the first-order maxima for these lines?

The straightforward way to find this separation is to repeat this calculation for $\lambda = 5895.9$ A and to subtract the two angles. A difficulty, which can best be appreciated by carrying out the calculation, is that we must carry a large number of significant figures to obtain a meaningful value for the difference between the angles. To calculate the difference in angular positions *directly*, let us write down Eq. 45-1, solved for $\sin\theta$, and differentiate it, treating θ and λ as variables:

$$\sin\theta = \frac{m\lambda}{d}$$

$$\cos\theta\, d\theta = \frac{m}{d}\, d\lambda.$$

If the wavelengths are close enough together, as in this case, $d\lambda$ can be replaced by $\Delta\lambda$, the actual wavelength difference; $d\theta$ then becomes $\Delta\theta$, the quantity we seek. This gives

$$\Delta\theta = \frac{m\,\Delta\lambda}{d\cos\theta} = \frac{(1)(5.9\text{ A})}{(25{,}400\text{ A})(\cos 13.3^\circ)} = 2.4 \times 10^{-4}\text{ radian} = 0.014^\circ.$$

Note that although the wavelengths involve five significant figures our calculation, done this way, involves only two or three, with consequent reduction in numerical manipulation. ◄

The quantity $d\theta/d\lambda$, called the *dispersion D* of a grating, is a measure of the angular separation produced between two incident monochromatic waves whose wavelengths differ by a small wavelength interval. From this example we see that

$$D = \frac{d\theta}{d\lambda} = \frac{m}{d\cos\theta}. \tag{45-4}$$

45–4 Resolving Power of a Grating

To distinguish light waves whose wavelengths are close together, the principal maxima of these wavelengths formed by the grating should be as narrow as possible. Expressed otherwise, the grating should have a high *resolving power R*, defined from

$$R = \frac{\lambda}{\Delta\lambda}. \tag{45-5}$$

Here λ is the mean wavelength of two spectrum lines that can barely be recognized as separate and $\Delta\lambda$ is the wavelength difference between them. The smaller $\Delta\lambda$ is, the closer the lines can be and still be resolved; hence the greater the resolving power R of the grating. It is to achieve a high resolving power that gratings with many rulings are constructed.

The resolving power of a grating is usually determined by the same consideration (that is, the Rayleigh criterion) that we used in Section 44–5 to determine the resolving power of a lens. If two principal maxima are to be barely resolved, they must, according to this criterion, have an angular separation $\Delta\theta$ such that the maximum of one line coincides with the first minimum of the other; see Fig. 44–11. If we apply this criterion, we can show that

$$R = Nm, \tag{45-6}$$

where N is the total number of rulings in the grating and m is the order. As expected, the resolving power is zero for the central principal maximum ($m = 0$), all wavelengths being undeflected in this order.

Let us derive Eq. 45–6. The angular separation between two principal maxima whose wavelengths differ by $\Delta\lambda$ is found from Eq. 45–4, which we recast as

$$\Delta\theta = \frac{m\,\Delta\lambda}{d\cos\theta}. \tag{45-4}$$

The Rayleigh criterion (Section 44–5) requires that this be equal to the angular separation between a principal maximum and its adjacent minimum. This is given from Eq. 45-3, dropping the subscript m in $\cos\theta_m$, as

$$\Delta\theta_m = \frac{\lambda}{dN\cos\theta}. \tag{45-3}$$

Equating Eqs. 45-4 and 45-3 leads to

$$R\ (= \lambda/\Delta\lambda) = Nm,$$

which is the desired relation.

▶ **Example 3.** In Example 2 how many rulings must a grating have if it is barely to resolve the sodium doublet in the third order?

From Eq. 45-5 the required resolving power is

$$R = \frac{\lambda}{\Delta\lambda} = \frac{5890\text{ A}}{(5895.9 - 5890.0)\text{A}} = 1000.$$

From Eq. 45-6 the number of rulings needed is

$$N = \frac{R}{m} = \frac{1000}{3} = 330.$$

This is a modest requirement. ◀

The resolving power of a grating must not be confused with its dispersion. Table 45-1 shows the characteristics of three gratings, each illuminated with light of $\lambda = 5890$ A, the diffracted light being viewed in the first order ($m = 1$ in Eq. 45-1).

Table 45-1

SOME CHARACTERISTICS OF THREE GRATINGS
($\lambda = 5890$ A, $m = 1$)

Grating	N	d, A	θ	R	D 10^{-3} degrees/A
A	10,000	25,400	13.3°	10,000	2.32
B	20,000	25,400	13.3°	20,000	2.32
C	10,000	13,700	25.5°	10,000	4.64

The student should verify that the values of D and R given in the table can be calculated from Eqs. 45-4 and 45-6, respectively.

For the conditions of use noted in Table 45-1, gratings A and B have the same *dispersion* and A and C have the same *resolving power*. Figure 45-9 shows the intensity patterns that would be produced by these gratings for two incident waves of wavelengths λ_1 and λ_2, in the vicinity of $\lambda = 5890$ A. Grating B, which has high resolving power, has narrow intensity maxima and is inherently capable of distinguishing lines that are much closer together in wavelength than those of Fig. 45-9. Grating C, which has high dispersion, produces twice the angular separation between rays λ_1 and λ_2 that grating B does.

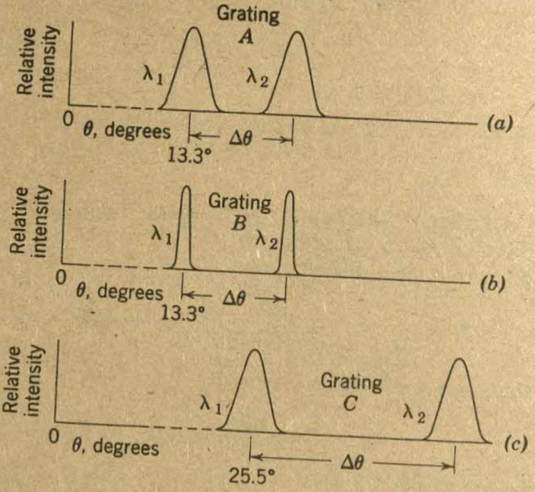


Fig. 45–9 The intensity patterns for light of wavelengths λ_1 and λ_2 near 5890 A, incident on the gratings of Table 45–1. Grating *B* has the highest resolving power and grating *C* the highest dispersion.

▶ **Example 4.** The grating of Example 1 has 8000 lines illuminated by light from a mercury vapor discharge. (*a*) What is the expected dispersion, in the third order, in the vicinity of the intense green line ($\lambda = 5460$ A)? Noting that $d = 31{,}700$ A, we have, from Eq. 45–1,

$$\theta = \sin^{-1}\frac{m\lambda}{d} = \sin^{-1}\frac{(3)(5460\text{ A})}{31{,}700\text{ A}} = \sin^{-1}0.517 = 31.1°.$$

From Eq. 45–4 we have

$$D = \frac{m}{d\cos\theta} = \frac{3}{(31{,}700\text{ A})(\cos 31.1°)} = 1.1 \times 10^{-4}\text{ radian/A} = 6.3 \times 10^{-3}\text{ deg/A}.$$

(*b*) What is the expected resolving power of this grating in the fifth order? Equation 45–6 gives

$$R = Nm = (8000)(5) = 40{,}000.$$

Thus near $\lambda = 5460$ A a wavelength difference $\Delta\lambda$ given by Eq. 45–5, or

$$\Delta\lambda = \frac{\lambda}{R} = \frac{5460\text{ A}}{40{,}000} = 0.14\text{ A},$$

can be distinguished. ◀

45–5 X-ray Diffraction

Figure 45–10 shows how X-rays are produced when electrons from a heated filament *F* are accelerated by a potential difference *V* and strike a metal target *T*. X-rays are electromagnetic radiation with wavelengths of the order of 1 A. This value is to be compared to 5500 A for the center of the visible spectrum. For such small wavelengths a standard optical diffraction grating, as normally employed, cannot be used. For $\lambda = 1$ A and $d = 30{,}000$ A,

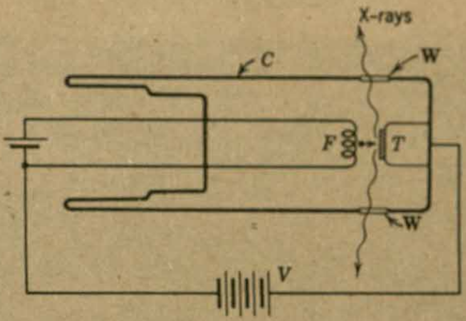


Fig. 45-10 X-rays are generated when electrons from heated filament F, accelerated by potential difference V, are brought to rest on striking metallic target T. W is a "window"—transparent to X-rays—in the evacuated metal container C.

for example, Eq. 45-1 shows that the first-order maximum occurs at

$$\theta = \sin^{-1}\frac{m\lambda}{d} = \sin^{-1}\frac{(1)(1\text{ A})}{3\times 10^4\text{ A}} = \sin^{-1} 0.33\times 10^{-4} = 0.002°.$$

This is too close to the central maximum to be practical. A grating with $d \cong \lambda$ is desirable, but, since X-ray wavelengths are about equal to atomic diameters, such gratings cannot be constructed mechanically.

In 1912 it occurred to the German physicist Max von Laue that a crystalline solid, consisting as it does of a regular array of atoms, might form a natural three-dimensional "diffraction grating" for X-rays. Figure 45-11 shows that if a collimated beam of X-rays, continuously distributed in wavelength, is allowed to fall on a crystal, such as sodium chloride, intense beams corresponding to constructive interference from the many diffracting centers

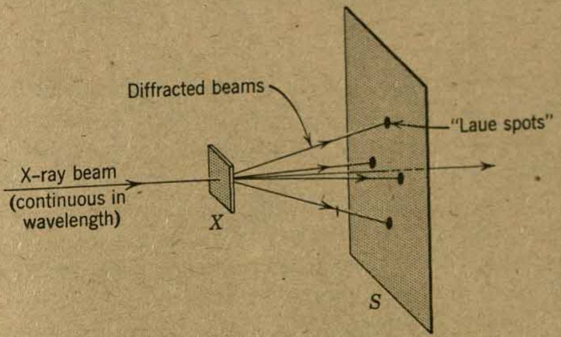


Fig. 45-11 A nonmonochromatic beam of X-rays falls on a crystal X, which may be NaCl. Strong diffracted beams appear in certain directions, forming a so-called Laue pattern on a photographic film S.

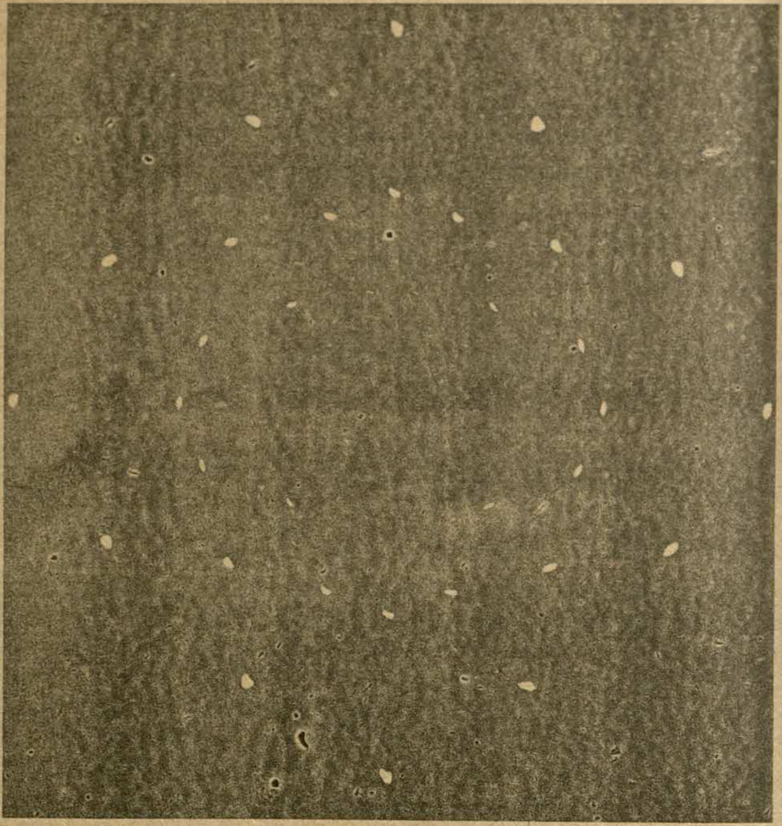


Fig. 45–12 Laue X-ray diffraction pattern from sodium chloride. A crystal of ordinary table salt was used in making this plate. (Courtesy of W. Arrington and J. L. Katz, X-ray Laboratory, Rensselaer Polytechnic Institute.)

of which the crystal is made up appear in certain sharply defined directions. If these beams fall on a photographic film, they form an assembly of "Laue spots." Figure 45–12, which is an actual example of these spots, shows that the hypothesis of Laue is indeed correct. The atomic arrangements in the crystal can be deduced from a careful study of the positions and intensities of the Laue spots * in much the same way that we might deduce the structure of an optical grating (that is, the detailed profile of its slits) by a study of the positions and intensities of the lines in the interference pattern.

Figure 45–13 shows how sodium and chlorine atoms (strictly, Na^+ and Cl^- ions) are stacked to form a crystal of sodium chloride. This pattern, which has *cubic* symmetry, is one of the many atomic arrangements exhibited by solids. The model represents the *unit cell for sodium chloride.* This is the smallest unit from which the crystal may be built up by repetition in three dimensions. The student should verify that no smaller assembly of atoms possesses this property. For sodium chloride the length of the cube edge of the unit cell is 5.62737 A.

---

* Other experimental arrangements have supplanted the Laue technique to a considerable extent today; the principle remains unchanged, however.

Each unit cell in sodium chloride has four sodium ions and four chlorine ions associated with it. In Fig. 45-13 the sodium ion in the center belongs entirely to the cell shown. Each of the other twelve sodium ions shown is shared with three adjacent unit cells so that each contributes one-fourth of an ion to the cell under consideration. The total number of sodium ions is then $1 + \frac{1}{4}(12) = 4$. By similar reasoning the student can show that although there are fourteen chlorine ions in Fig. 45-13 only four are associated with the unit cell shown.

The unit cell is the fundamental repetitive diffracting unit in the crystal, corresponding to the slit (and its adjacent opaque strip) in the optical dif-

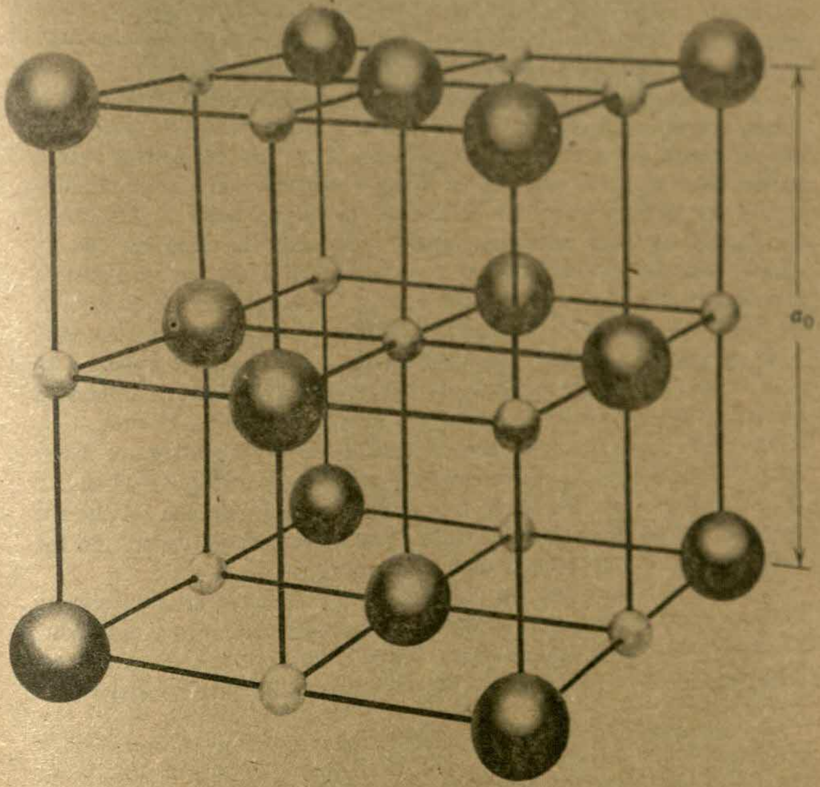
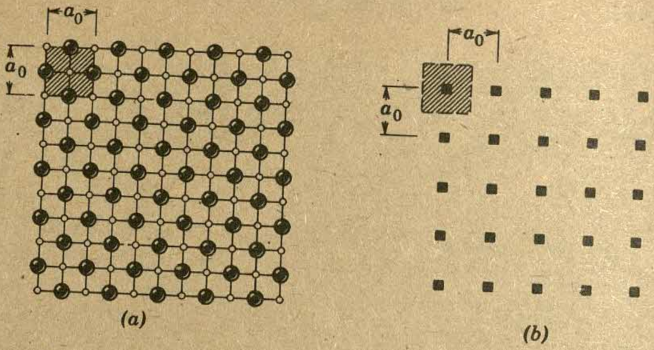


Fig. 45-13 A model showing how Na^+ and Cl^- ions are stacked to form a unit cell of NaCl. The small spheres represent sodium ions, the large ones chlorine. The edge a_0 of the (cubical) unit cell is 5.62737 A.



(a)

(b)

Fig. 45–14 (a) A section through a crystal of sodium chloride, showing the sodium and chlorine ions. (b) The corresponding unit cells in this section, each cell being represented by a small black square.

fraction grating of Fig. 45–1. Figure 45–14*a* shows a particular plane in a sodium chloride crystal. If each unit cell intersected by this plane is represented by a small cube, Fig. 45–14*b* results. The student may imagine each of these figures extended indefinitely in three dimensions.

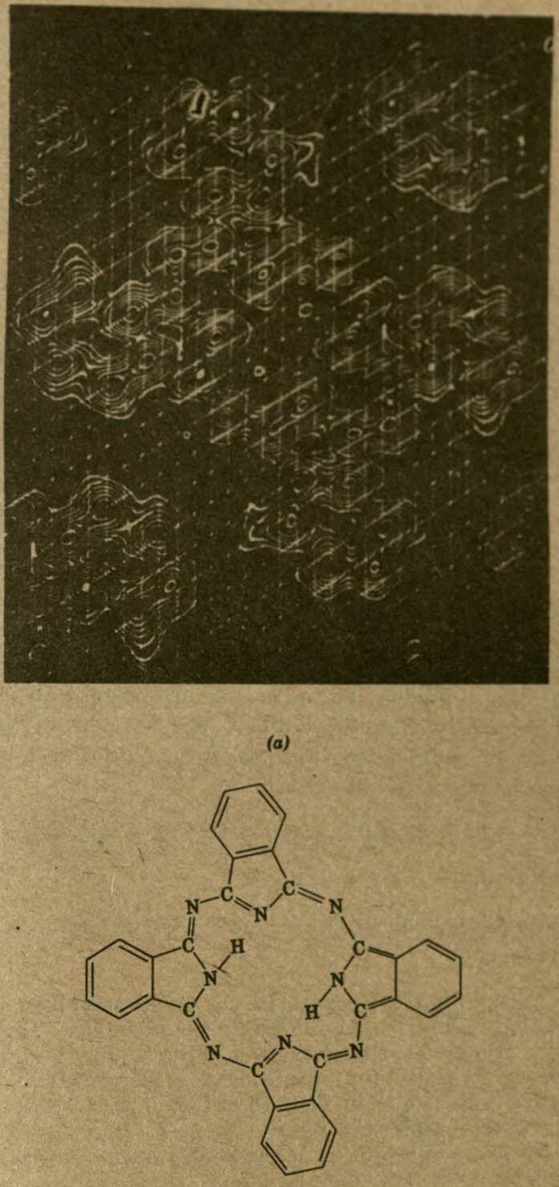
Let us treat each small cube in Fig. 45–14*b* as an elementary diffracting center, corresponding to a slit in an optical grating. The *directions* (but not the intensities) of all the diffracted X-ray beams that can emerge from a sodium chloride crystal (for a given X-ray wavelength and a given orientation of the incident beam) are determined by the geometry of this three-dimensional lattice of diffracting centers. In exactly the same way the *directions* (but not the intensities) of all the diffracted beams that can emerge from a particular optical grating (for a given wavelength and orientation of the incident beam) are determined only by the geometry of the grating, that is, by the grating spacing *d*. Representing the unit cell by what is essentially a point, as in Fig. 45–14*b*, corresponds to representing the slits in a diffraction grating by lines, as we did in discussing Young's experiment in Section 43–1.

The *intensities* of the lines from an optical diffraction grating are determined by the diffracting characteristics of a single slit, as Fig. 44–14 shows. In the idealized case of Fig. 45–1 these characteristics depend on the slit width *a*. In practical optical gratings these characteristics depend on the detailed shape of the profile of the grating rulings.

In exactly the same way the *intensities* of the diffracted beams emerging from a crystal depend on the diffracting characteristics of the unit cell.* Fundamentally the X-rays are diffracted by electrons, diffraction by nuclei being negligible in most cases. Thus the diffracting characteristics of a

---

* For some directions in which a beam might be expected to emerge, from interference considerations, no beam will be found because the diffracting characteristics of the unit cell are such that no energy is diffracted in that direction. Similarly, in optical gratings some lines, permitted by interference considerations, may not appear if their predicted positions coincide with a null in the single-slit diffraction pattern (see Fig. 44–12).



(a)

(b)

Fig. 45–15 (*a*) A photograph of an oscilloscope screen arranged to display projected electron density contours for phthalocyanine ($C_{32}H_{18}N_8$). Such plots, constructed electronically from X-ray diffraction data by an analog computer, provide a vivid picture of the structure of molecules. (Courtesy of Ray Pepinsky.) (*b*) A structural representation of the molecule phthalocyanine. The student should make a detailed comparison with (*a*), locating the various atoms identified in (*b*). Note that hydrogen atoms, which contain only a single electron, are not prominent in (*a*).

unit cell depend on how the electrons are distributed throughout the volume of the cell By studying the *directions* of diffracted X-ray beams, we can learn the basic symmetry of the crystal. By also studying the *intensities* we can learn how electrons are distributed in the unit cell. Figure 45–15 shows the average density of electrons projected onto a particular plane through a unit cell of a crystal of phthalocyanine. This remarkable figure suggests something of the full power of X-ray methods for studying the structure of solids.

45–6 Bragg's Law

Bragg's law predicts the conditions under which diffracted X-ray beams from a crystal are possible. In deriving it, we ignore the structure of the unit cell, which is related only to the intensities of these beams. The dashed sloping lines in Fig. 45–16*a* represent the intersection with the plane of the figure of an arbitrary set of planes passing through the elementary diffracting centers. The perpendicular distance between adjacent planes is d. Many other such families of planes, with different *interplanar spacings*, can be defined.

Figure 45–16*b* shows a plane wave that lies in the plane of the figure falling on one member of the family of planes defined in Fig. 45–16*a*, the incident rays making an angle θ with the plane.* Consider a family of diffracted rays lying in the plane of Fig. 45–16*b* and making an angle β with the plane containing the elementary diffracting centers. The diffracted rays will combine to produce maximum intensity if the path difference between adjacent rays is an integral number of wavelengths or

$$ae - bd = h(\cos\beta - \cos\theta) = l\lambda \qquad l = 0, 1, 2, \ldots \qquad (45\text{–}7)$$

For $l = 0$ this leads to

$$\beta = \theta,$$

and the plane of atoms acts like a mirror for the incident wave, no matter what the value of θ.

For other values of l, β does not equal θ, but the diffracted beam can always be regarded as being "reflected" from a *different* set of planes than that shown in Fig. 45–16*a* with a different interplanar spacing d. Since we wish to describe each diffracted beam as a "reflection" from a particular set of planes and since we are dealing in the present argument only with the particular set of planes shown in Fig. 45–16*a* we ignore all values of l other than $l = 0$ in Eq. 45–7. It can also be shown that a plane of diffracting centers acts like a mirror (that is, $\beta = \theta$) whether or not the incident wave lies in the plane of Fig. 45–16*b*.

Figure 45–16*c* shows an incident wave striking the *family* of planes, a single member of which was considered in Fig. 45–16*b*. For a single plane, mirrorlike "reflection" occurs for *any* value of θ, as we have seen. To have a constructive interference in the beam diffracted from the entire family of planes

---

[*] In X-ray diffraction it is customary to specify the direction of a wave by giving the angle between the ray and the plane (the *glancing angle*) rather than the angle between the ray and the normal.

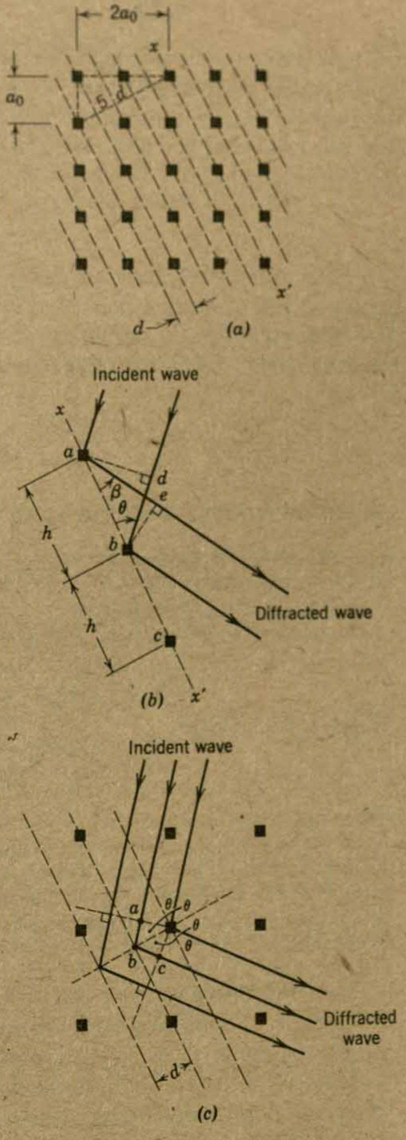


Fig. 45–16 (*a*) A section through the NaCl unit cell lattice of Fig. 45–14*b*. The dashed sloping lines represent an arbitrary family of planes, with interplanar spacing d (*b*) An incident wave falls, at grazing angle θ, on one of the planes, xx', shown in (*a*). (*c*) An incident wave falls on the entire family of planes shown in (*a*). A strong diffracted wave is formed.

in the direction θ, the rays from the separate planes must reinforce each other. This means that the path difference for rays from adjacent planes (*abc* in Fig. 45–16*c*) must be an integral number of wavelengths or

$$2d \sin \theta = m\lambda \qquad m = 1, 2, 3, \ldots \tag{45–8}$$

This relation is called *Bragg's law* after W. L. Bragg who first derived it. The quantity d in this equation (the interplanar spacing) is the perpendicular distance between the planes. For the planes of Fig. 45–16*a* analysis shows that d is related to the unit cell dimension a_0 by

$$d = \frac{a_0}{\sqrt{5}}. \tag{45–9}$$

If an incident *monochromatic* X-ray beam falls at an *arbitrary* angle θ on a particular set of atomic planes, a diffracted beam will *not* result because Eq. 45–8 will not, in general, be satisfied. If the incident X-rays are *continuous* in wavelength, diffracted beams will result when wavelengths given by

$$\lambda = \frac{2d \sin \theta}{m} \qquad m = 1, 2, 3, \ldots$$

are present in the incident beam (see Eq. 45–8).

▶ **Example 5.** At what angles must an X-ray beam with $\lambda = 1.10$ A fall on the family of planes represented in Fig. 45–16*c* if a diffracted beam is to exist? Assume the material to be sodium chloride.

The interplanar spacing d for these planes is given by Eq. 45–9 or

$$d = \frac{a_0}{\sqrt{5}} = \frac{5.63 \text{ A}}{2.24} = 2.52 \text{ A}.$$

Equation 45–8 gives

$$\sin \theta = \frac{m\lambda}{2d} = \frac{(m)(1.10 \text{ A})}{(2)(2.52 \text{ A})} = 0.218m.$$

Diffracted beams are possible at $\theta = 12.6°$ $(m = 1)$, $\theta = 25.9°$ $(m = 2)$, $\theta = 40.9°$ $(m = 3)$, and $\theta = 60.7°$ $(m = 4)$. Higher-order beams cannot exist because they require $\sin \theta$ to exceed unity. Actually, the odd-order beams $(m = 1, 3)$ prove to have zero intensity because the unit cell in cubic crystals such as NaCl has diffracting properties such that the intensity of the light scattered in these orders is zero (see Problem 26). ◀

X-ray diffraction is a powerful tool for studying the arrangements of atoms in crystals. To do so quantitatively requires that the wavelength of the X-rays be known. In one of several approaches to this problem the unit cell dimension for NaCl * is determined by a method that does not involve X-rays. X-ray diffraction measurements on NaCl can then be used to determine the wavelength of the X-ray beam which in turn can be used to determine the structures of solids other than NaCl.

---

* In practice, calcite ($CaCO_3$) proves to be more useful as a standard crystal for a number of technical reasons.

If ρ is the measured density of NaCl, we have, for the unit cell of Fig. 45–13,* recalling that each unit cell contains four NaCl "molecules,"

$$\rho = \frac{m}{V} = \frac{4m_{\mathrm{NaCl}}}{a_0^3}.$$

Here m_{NaCl}, the mass of a NaCl molecule, is given by

$$m_{\mathrm{NaCl}} = \frac{M}{N_0},$$

where M is the molecular weight of NaCl and N_0 is Avogadro's number. Combining these two equations and solving for a_0 yields

$$a_0 = \left(\frac{4M}{N_0\rho}\right)^{1/3}$$

which permits us to calculate a_0. Once a_0 is known, the wavelengths of monochromatic X-ray beams can be found, using Bragg's law (Eq. 45–8).

QUESTIONS

1. Discuss this statement: "A diffraction grating can just as well be called an interference grating."

2. For the simple spectroscope of Fig. 45–7, show (*a*) that θ increases with λ for a grating and (*b*) that θ decreases with λ for a prism.

3. You are given a photograph of a spectrum on which the angular positions and the wavelengths of the spectrum lines are marked. (*a*) How can you tell whether the spectrum was taken with a prism or a grating instrument? (*b*) What information could you gather about either the prism or the grating from studying such a spectrum?

4. Assume that the limits of the visible spectrum are 4300 and 6800 A. Is it possible to design a grating, assuming that the incident light falls normally on it, such that the first-order spectrum *barely overlaps* the second-order spectrum?

5. (*a*) Why does a diffraction grating have closely spaced rulings? (*b*) Why does it have a large number of rulings?

6. The relation $R = Nm$ suggests that the resolving power of a given grating can be made as large as desired by choosing an arbitrarily high order of diffraction. Discuss.

7. Show that at a given wavelength and a given angle of diffraction the resolving power of a grating depends only on its width W ($= Nd$).

8. According to Eq. 45–3 the principal maxima become wider (that is, $\Delta\theta_m$ increases) the higher the order m (that is, the larger θ_m becomes). According to Eq. 45–6 the resolving power becomes greater the higher the order m. Explain this apparent paradox.

9. Is the pattern of Fig. 45–12 more properly described as a diffraction pattern or as an interference pattern?

10. For a given family of planes in a crystal, can the wavelength of incident X-rays be (*a*) too large or (*b*) too small to form a diffracted beam?

11. If a parallel beam of X-rays of wavelength λ is allowed to fall on a randomly oriented crystal of any material, generally no intense diffracted beams will occur. Such beams

---

* This relation cannot be written down unless it is known that the structure of NaCl is cubic. This can be determined, however, by inspection of the symmetry of the spots in Fig. 45–12; the wavelength of the X-rays need not be known.

appear if (*a*) the X-ray beam consists of a continuous distribution of wavelengths rather than a single wavelength or (*b*) the specimen is not a single crystal but a finely divided powder. Explain.

12. Why cannot a simple cube of edge $a_0/2$ in Fig. 45–13 be used as a unit cell for sodium chloride?

13. How would you *measure* (*a*) the dispersion D and (*b*) the resolving power R for either a prism or a grating spectrograph.

PROBLEMS

1. Given a grating with 4000 lines/cm, how many orders of entire visible spectrum (4000–7000 A) can be produced?

2. Derive this expression for the intensity pattern for a three-slit "grating":

$$I_\theta = \tfrac{1}{9}I_m(1 + 4\cos\phi + 4\cos^2\phi),$$

where

$$\phi = \frac{2\pi d \sin\theta}{\lambda}.$$

Assume that $a \ll \lambda$ and be guided by the derivation of the corresponding double-slit formula (Eq. 43–9).

3. (*a*) Using the result of Problem 2, show that the half-width of the fringes for a three-slit diffraction pattern, assuming θ small enough so that $\sin\theta \cong \theta$, is

$$\Delta\theta \cong \frac{\lambda}{3.2d}.$$

(*b*) Compare this with the expression derived for the two-slit pattern in Problem 10, Chapter 43. (*c*) Do these results support the conclusion that for a fixed slit spacing the interference maxima become sharper as the number of slits is increased?

4. Using the result of Problem 2, show that a three-slit "grating" has only one secondary maximum. Find its location and its relative intensity.

5. A grating designed for use in the infrared region of the electromagnetic spectrum is "blazed" to concentrate all its intensity in the first order ($m = 1$) for $\lambda = 80{,}000$ A. If visible light ($4000 \text{ A} < \lambda < 7000 \text{ A}$) were allowed to fall on this grating, what visual appearance would the diffracted beams present?

6. The central intensity maximum formed by a grating, along with its subsidiary secondary maxima, can be viewed as the diffraction pattern of a single "slit" whose width is that of the entire grating. Treating the grating as a single wide slit, assuming that $m = 0$, and using the methods of Section 44–4, show that Eq. 45–2 can be derived.

7. A grating has 8000 rulings/in. For what wavelengths in the visible spectrum can fifth-order diffraction be observed?

8. A diffraction grating has 5000 rulings/in., and a strong diffracted beam is noted at $\theta = 30°$. (*a*) What are the possible wavelengths of the incident light? (*b*) How could you identify them in an actual case?

9. A diffraction grating 2.0 cm wide has 6000 rulings. At what angles will maximum-intensity beams occur if the incident radiation has a wavelength of 5890 A?

10. Assume that the limits of the visible spectrum are arbitrarily chosen as 4300 and 6800 A. Design a grating that will spread the first-order spectrum through an angular range of 20°.

11. A grating has 8000 rulings/in. and is illuminated at normal incidence by white light. A spectrum is formed on a screen 30 cm from the grating. If a 1.0-cm square hole is cut in the screen, its inner edge being 5.0 cm from the central maximum, what range of wavelengths passes through the hole?

12. Assume that light is incident on a grating at an angle ψ as shown in Fig. 45–17. Show that the condition for a diffraction maximum is

$$d(\sin \psi + \sin \theta) = m\lambda \qquad m = 0, 1, 2, \ldots.$$

Only the special case $\psi = 0$ has been treated in this chapter (compare Eq. 45–1).

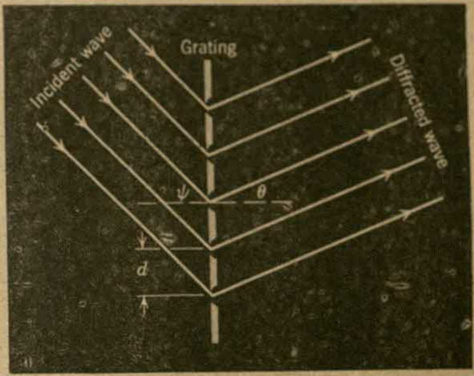


Fig. 45–17

13. Show that in a grating with alternately transparent and opaque strips of equal width all the even orders (except $m = 0$) are absent.

14. A transmission grating with $d = 1.50 \times 10^{-4}$ cm is illuminated at various angles of incidence by light of wavelength 6000 A. Plot as a function of angle of incidence (0 to 90°) the angular deviation of the first-order diffracted beam from the incident direction.

15. Derive Eq. 45–3, that is, the expression for $\Delta\theta_m$, the angular distance between a principal maximum of order m and either adjacent minimum.

16. A three-slit grating has separation d between adjacent slits. If the middle slit is covered up, will the half-width of the intensity maxima become broader or narrower? See Problem 3 and also Problem 10, Chapter 43.

17. A grating has 40,000 rulings spread over 3.0 in. (*a*) What is its expected dispersion D for sodium light ($\lambda = 5890$ A) in the first three orders? (*b*) What is its resolving power in these orders?

18. In a particular grating the sodium doublet (see Example 2) is viewed in third order at 80° to the normal and is barely resolved. Find (*a*) the grating spacing and (*b*) the total width of the rulings.

19. A source containing a mixture of hydrogen and deuterium atoms emits a red doublet at $\lambda = 6563$ A whose separation is 1.8 A. Find the minimum number of lines needed in a diffraction grating which can resolve these lines in the first order.

20. Show that the dispersion of a grating can be written as

$$D = \frac{\tan \theta}{\lambda}.$$

21. A grating has 6000 rulings/cm and is 6.0 cm wide. (*a*) What is the smallest wavelength interval that can be resolved in the third order at $\lambda = 5000$ A? (*b*) For this wavelength and this grating, can the resolution be improved? How?

22. Light containing a mixture of two wavelengths, 5000 A and 6000 A, is incident normally on a diffraction grating. It is desired (1) that the first and second principal maxima for each wavelength appear at $\theta \leq 30°$, (2) that the dispersion be as high as possible, *and* (3) that the third order for 6000 A be a missing order. (*a*) What is the separation

between adjacent slits? (*b*) What is the smallest possible individual slit width? (*c*) Name *all* orders for 6000 A that actually appear on the screen with the values chosen in (*a*) and (*b*).

23. Light of wavelength 6000 A is incident normally on a diffraction grating. Two *adjacent* principal maxima occur at $\sin\theta = 0.2$ and $\sin\theta = 0.3$, respectively. The fourth order is a missing order. (*a*) What is the separation between adjacent slits? (*b*) What is the smallest possible individual slit width? (*c*) Name *all* orders actually appearing on the screen with the values chosen in (*a*) and (*b*).

24. An optical grating with a spacing $d = 15{,}000$ A is used to analyze soft X-rays of wavelength $\lambda = 5.0$ A. The angle of incidence θ is $90° - \gamma$, where γ is a *small* angle. The first order maximum is found at an angle $\theta = 90° - 2\beta$. Find the value of β.

25. Monochromatic X-rays ($\lambda = 1.20$ A) fall on a crystal of sodium chloride, making an angle of 45° with a reference line as shown in Fig. 45–18. Through what angles must the crystal be turned to give a diffracted beam associated with the planes shown? Assume that the crystal is turned about an axis that is perpendicular to the plane of the page. Ignore the possibility (see Problem 26) that some of these beams may be of zero intensity.

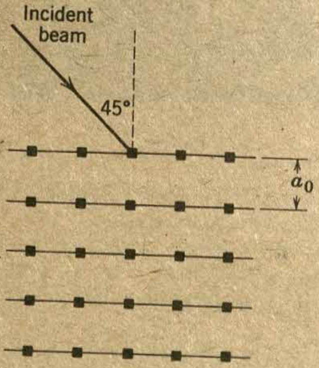


Fig. 45–18

26. *Missing orders in X-ray diffraction.* In Problem 25 the $m = 1$ beam, permitted by interference considerations, has zero intensity because of the diffracting properties of the unit cell for this geometry of beams and crystal. Prove this. (Hint: Show that the "reflection" from an atomic plane through the top of a layer of unit cells is canceled by a "reflection" from a plane through the middle of this layer of cells. All odd-order beams prove to have zero intensity.)

27. Assume that the incident X-ray beam in Fig. 45–18 is not monochromatic but contains wavelengths in a band from 0.95 to 1.30 A. Will diffracted beams, associated with the planes shown, occur? Assume $a_0 = 2.75$ A.

28. In comparing the wavelengths of two monochromatic X-ray lines, it is noted that line *A* gives a first-order reflection maximum at a glancing angle of 30° to the smooth face of a crystal. Line *B*, known to have a wavelength of 0.97 angstroms, gives a third-order reflection maximum at an angle of 60° from the same face of the same crystal. Find the wavelength of line *A*.

Polarization

CHAPTER 46

46–1 Polarization

Light, like all electromagnetic radiation, is predicted by electromagnetic theory to be a *transverse wave,* the directions of the vibrating electric and magnetic vectors being at right angles to the direction of propagation instead of parallel to it as in a longitudinal wave. The transverse waves of Figs. 46–1 and 39–11 have the additional characteristic that they are *plane-polarized.* This means that the vibrations of the **E** vector are parallel to each other for all points in the wave. At any such point the vibrating **E** vector and the direction of propagation form a plane, called the *plane of vibration;* in a plane-polarized wave all such planes are parallel.

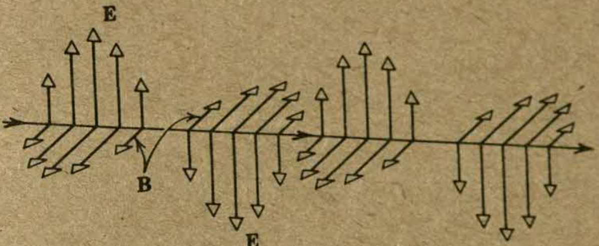


Fig. 46–1 An instantaneous "snapshot" of a plane-polarized wave showing the vectors **E** and **B** along a particular ray. The wave is moving to the right with speed c. The plane containing the vibrating **E** vector and the direction of propagation is a *plane of vibration.*

The transverse nature of light waves cannot be deduced from the interference or diffraction experiments so far described because longitudinal waves such as sound waves also show these effects. An experimental basis for believing that light waves are transverse was provided by Thomas Young in 1817. Two of his contemporaries, Dominique-François Arago (1786–1853) and Augustin Jean Fresnel (1788–1827), were able, by allowing a light beam to fall on a crystal of calcite, to produce two separate beams (see Section 46–4). Astonishingly, these beams, although coherent, produced no interference fringes but only a uniform illumination. Young deduced from this that light must be a transverse wave and that the planes of vibration in the two beams must be at right angles to each other. Wave disturbances that act at right angles to each other cannot show interference effects; the student is asked to prove this in Problem 9. Young's words to Arago were these:

> I have been reflecting on the possibility of giving an imperfect explanation of the affection of light which constitutes polarization without departing from the genuine doctrine of undulations. It is a principle in this theory that all undulations are simply propagated through homogeneous mediums in concentric spherical surfaces like the undulations of sound, consisting simply in the direct and retrograde motions of the particles in the direction of the radius with their concomitant condensation and rarefactions [that is, longitudinal waves]. And yet, it is possible to explain in this theory a transverse vibration, propagated also in the direction of the radius, and with equal velocity, the motions of the particles being in a certain constant direction with respect to that radius; and this is a *polarization*.

Note how Young presents the possibility of a transverse vibration as a novel idea, light having been generally—but incorrectly—assumed to be a longitudinal vibration.

In a plane-polarized transverse wave it is necessary to specify two directions, that of the wave disturbance (**E**, say) and that of propagation. In a longitudinal wave these directions are identical. In plane-polarized transverse waves, but not in longitudinal waves, we may thus expect a lack of symmetry about the direction of propagation. Electromagnetic waves in the radio and microwave range exhibit this lack of symmetry readily. Such a wave, generated by the surging of charge up and down in the dipole that forms the transmitting antenna of Fig. 46–2, has (at large distances from the dipole and at right angles to it) an electric field vector parallel to the dipole axis. When this plane-polarized wave falls on a second dipole connected to a microwave detector, the alternating electric component of the wave will cause electrons to surge back and forth in the receiving antenna, producing a reading on the detector. If we turn the receiving antenna through 90° about the direction of propagation, the detector reading drops to zero. In this orientation the electric field vector is not able to cause charge to move along the dipole axis because it points at right angles to this axis. We can reproduce the experiment of Fig. 46–2 by turning the receiving antenna of a television set (assumed an electric dipole type) through 90° about an axis that points toward the transmitting station.

Common sources of visible light differ from radio and microwave sources

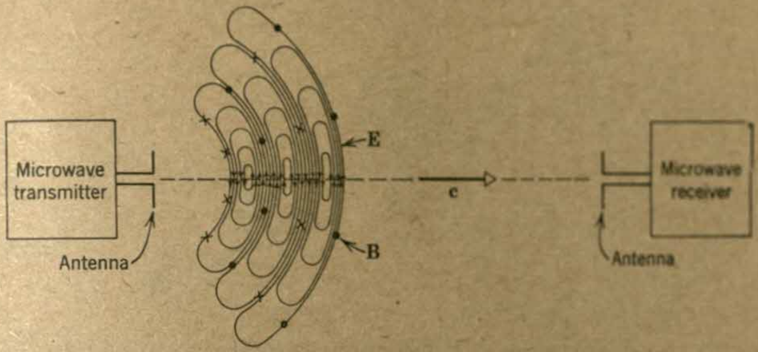


Fig. 46-2 The vectors **E** in the transmitted wave are parallel to the axis of the receiving antenna so that the wave will be detected. If the receiving antenna is rotated through 90° about the direction of propagation, no signal will be detected.

in that the elementary radiators, that is, the atoms and molecules, act independently. The light propagated in a given direction consists of independent wavetrains whose planes of vibration are randomly oriented about the direction of propagation, as in Fig. 46-3b. Such light, though still transverse, is *unpolarized*. The random orientation of the planes of vibration produces symmetry about the propagation direction, which, on casual study, conceals the true transverse nature of the waves. To study this transverse nature, a way must be found to unsort the different planes of vibration.

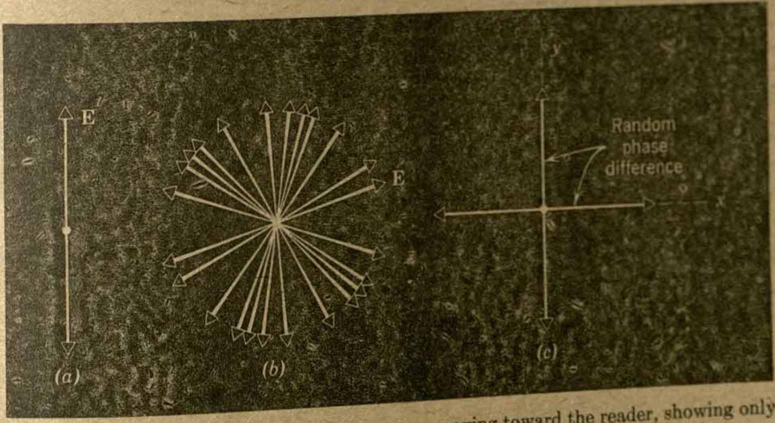


Fig. 46-3 (a) A plane-polarized transverse wave moving toward the reader, showing only the electric vector. (b) An unpolarized transverse wave viewed as a random superposition of many plane-polarized wavetrains. (c) A second, completely equivalent, description of an unpolarized transverse wave; here the unpolarized wave is viewed as two plane-polarized waves *with a random phase difference*. The orientation of the x and y axes about the propagation direction is completely arbitrary.

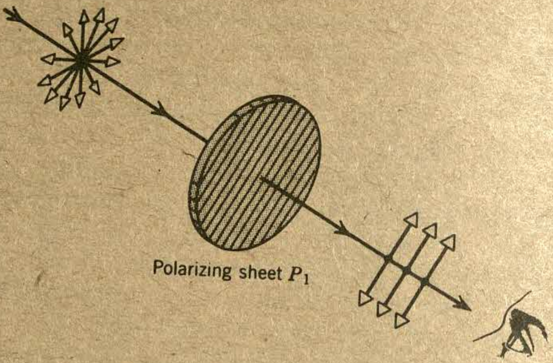


Fig. 46–4 A polarizing sheet produces plane-polarized light from unpolarized light. The parallel lines, which are not actually visible on the sheet, suggest the characteristic polarizing direction of the sheet.

46–2 Polarizing Sheets

Figure 46–4 shows unpolarized light falling on a sheet of commercial polarizing material called *Polaroid.** There exists in the sheet a certain characteristic polarizing direction, shown by the parallel lines. *The sheet will transmit only those wave-train components whose electric vectors vibrate parallel to this direction and will absorb those that vibrate at right angles to this direction.* The emerging light will be plane-polarized. This polarizing direction is established during the manufacturing process by embedding certain long-chain molecules in a flexible plastic sheet and then stretching the sheet so that the molecules are aligned parallel to each other. Polarizing sheets 2 ft wide and 100 ft long may be produced.

In Fig. 46–5 the polarizing sheet or *polarizer* lies in the plane of the page and the direction of propagation is into the page. The arrow **E** shows the plane of vibration of a randomly selected wavetrain falling on the sheet. Two vector components, $\mathbf{E}_x$ (of magnitude $E \sin \theta$) and $\mathbf{E}_y$ (of magnitude $E \cos \theta$), can replace **E**, one parallel to the polarizing direction and one at right angles to it. Only the former will be transmitted; the other is absorbed within the sheet.

Let us place a second polarizing sheet P_2 (usually called, when so used, an *analyzer*) as in Fig. 46–6. If P_2 is rotated about the direction of propagation, there are two positions, 180° apart, at which the transmitted light intensity is almost zero; these are the positions in which the polarizing directions of P_1 and P_2 are at right angles.

If the amplitude of the plane-polarized light falling on P_2 is E_m, the amplitude of the light that emerges is $E_m \cos \theta$, where θ is the angle between the polarizing directions of P_1 and P_2. Recalling that the intensity of the light

---

* There are other ways of producing polarized light without using this well-known commercial product. We mention some of them below.

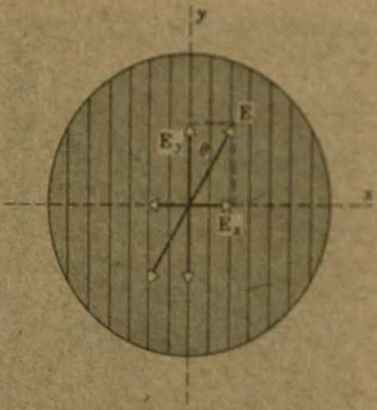


Fig. 46–5 A wavetrain **E** is equivalent to two component wavetrains $\mathbf{E}_y$ and $\mathbf{E}_x$. Only the former is transmitted by the polarizer.

beam is proportional to the square of the amplitude, we see that the transmitted intensity I varies with θ according to

$$I = I_m \cos^2 \theta, \tag{46–1}$$

in which I_m is the maximum value of the transmitted intensity. It occurs when the polarizing directions of P_1 and P_2 are parallel, that is, when $\theta = 0$ or 180°. Figure 46–7*a*, in which two overlapping polarizing sheets are in the parallel position ($\theta = 0$ or 180° in Eq. 46–1) shows that the light transmitted through the region of overlap has its maximum value. In Fig. 46–7*b* one or the other of the sheets has been rotated through 90° so that θ in Eq. 46–1 has the value 90 or 270°; the light transmitted through the region of overlap is now a minimum.

Equation 46–1, called the law of Malus, was discovered by Étienne Louis Malus (1775–1812) experimentally in 1809, using polarizing techniques other

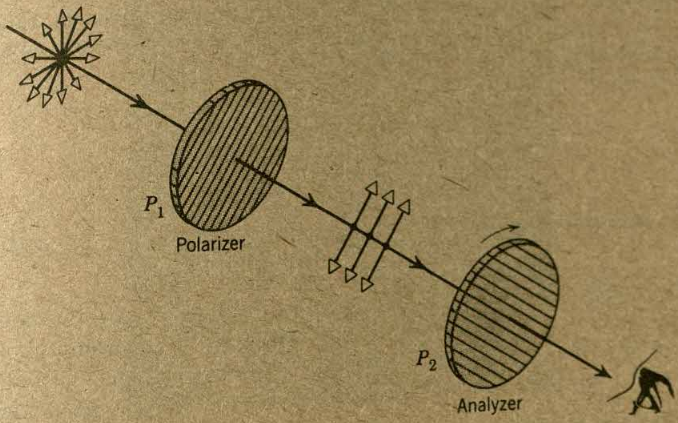


Fig. 46–6 Unpolarized light is not transmitted by crossed polarizing sheets.

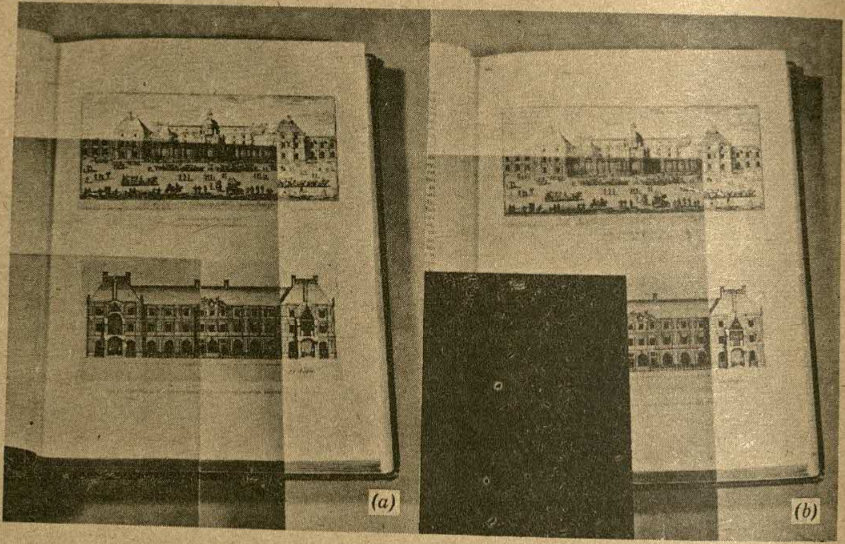


Fig. 46–7 Two square sheets of Polaroid are laid over a book. In (*a*) the axes of polarization of the two sheets are parallel and light passes through both sheets. In (*b*) one sheet has been rotated 90° and no light passes through. The book is opened to an illustration of the Luxembourg Palace in Paris. Malus discovered the phenomenon of polarization by reflection while looking at sunlight reflected off the palace windows through a calcite crystal.

than those so far described (see Section 46–3). Equation 46–1 describes precisely the lack of symmetry about the propagation direction that must be exhibited by plane-polarized transverse waves. Longitudinal waves could not possibly show such effects. Interestingly enough the human eye, under certain conditions, can detect polarized light.*

▶ **Example 1.** Two polarizing sheets have their polarizing directions parallel so that the intensity I_m of the transmitted light is a maximum. Through what angle must either sheet be turned if the intensity is to drop by one-half?

From Eq. 46–1, since $I = \frac{1}{2}I_m$, we have

$$\tfrac{1}{2}I_m = I_m \cos^2\theta$$

or

$$\theta = \cos^{-1} \pm \frac{1}{\sqrt{2}} = \pm 45°, \pm 135°.$$

The same effect is obtained no matter which sheet is rotated or in which direction. ◀

Historically polarization studies were made to investigate the nature of light. Today we reverse the procedure and deduce something about the nature of an object from the polarization state of the light emitted by or scattered from that object. It has been possible to deduce, from studies of the polarization of light reflected from them, that the grains of cosmic dust present in our galaxy have been oriented in the weak galactic magnetic field ($\sim 2 \times 10^{-4}$ gauss) so that their long dimension is parallel to this field. Polarization studies have shown that Saturn's rings consist of ice crystals. The size and shape of virus particles can be determined by the polarization of ultraviolet light scattered from them. Much useful information about the structure of atoms and nuclei is gained from polarization studies of their emitted radiations in all parts of the electromagnetic spectrum. Thus we have a useful research technique for structures ranging in size from a galaxy ($\sim 10^{+20}$ meters) to a nucleus ($\sim 10^{-14}$ meter). Polarized light also has many practical applications in industry and in engineering science.

---

* The so-called *Haidinger's brushes*; the interested student is referred to *Concepts of Classical Optics*, John Strong, W. H. Freeman & Co., 1958.

46–3 Polarization by Reflection

Malus discovered in 1809 that light can be partially or completely polarized by reflection. Anyone who has watched the sun's reflection in water, while wearing a pair of sunglasses made of polarizing sheet, has probably noticed the effect. It is necessary only to tilt the head from side to side, thus rotating the polarizing sheets, to observe that the intensity of the reflected sunlight passes through a minimum.

Figure 46–8 shows an unpolarized beam falling on a glass surface. The **E** vector for each wavetrain in the beam can be resolved into two components, one perpendicular to the plane of incidence—which is the plane of Fig. 46–8—and one lying in this plane. The first component, represented by the dots, is called the *σ-component*, from the German *senkrecht*, meaning perpendicular. The second component, represented by the arrows, is called the *π-component* (for *parallel*). On the average, for completely unpolarized incident light, these two components are of equal amplitude.

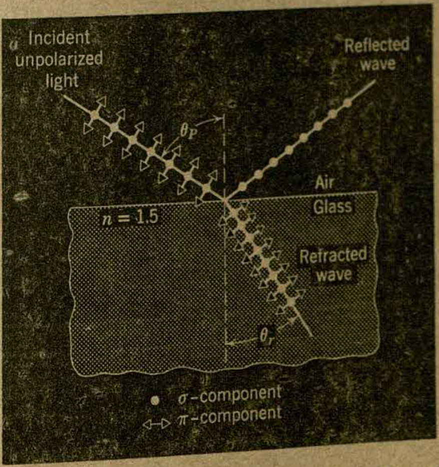


Fig. 46–8 For a particular angle of incidence θ_p, the reflected light is completely polarized, as shown. The transmitted light is partially polarized.

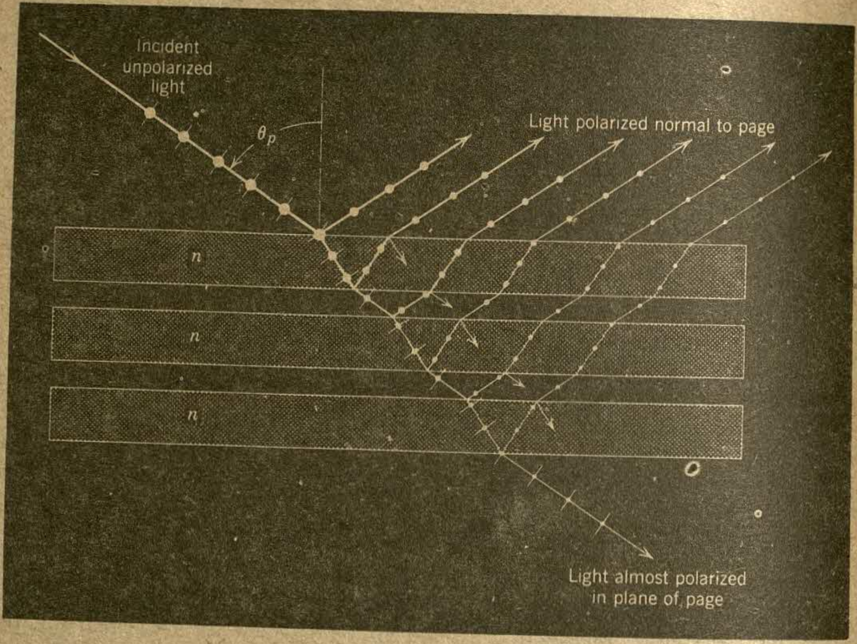


Fig. 46–9 Polarization of light by a stack of glass plates. Unpolarized light is incident on a stack of glass plates at Brewster's angle θ_p. (Polarization in the plane of the page is shown by the short lines and polarization normal to the page by the dots.) All light reflected out of the original ray is polarized normal to the page. After passing through several reflecting interfaces, the light in the original ray no longer contains any appreciable component polarized normal to the page.

Experimentally, for glass or other dielectric materials, there is a particular angle of incidence, called the *polarizing angle* θ_p, at which the reflection coefficient for the π-component is zero. This means that the beam reflected from the glass, although of low intensity, is plane-polarized, with its plane of vibration at right angles to the plane of incidence. This polarization of the reflected beam can easily be verified by analyzing it with a polarizing sheet.

The π-component at the polarizing angle is entirely refracted; the σ-component is only partially refracted. Thus the transmitted beam, which is of high intensity, is only partially polarized. By using a stack of glass plates rather than a single plate, reflections from successive surfaces occur and the intensity of the emerging reflected (σ-component) beam can be increased (see Fig. 46–9). By the same token, the σ-components are progressively removed from the transmitted beam, making it more completely π-polarized.

At the polarizing angle it is found experimentally that the reflected and the refracted beams are at right angles, or (Fig. 46–8)

$$\theta_p + \theta_r = 90°.$$

From Snell's law,

$$n_1 \sin \theta_p = n_2 \sin \theta_r.$$

Combining these equations leads to

$$n_1 \sin \theta_p = n_2 \sin (90^\circ - \theta_p) = n_2 \cos \theta_p$$

or

$$\tan \theta_p = \frac{n_2}{n_1}, \tag{46-2}$$

where the incident ray is in medium one and the refracted ray in medium two. This can be written as

$$\tan \theta_p = n, \tag{46-3}$$

where n $(= n_2/n_1)$ is the index of refraction of medium two with respect to medium one. Equation 46-3 is known as *Brewster's law* after Sir David Brewster (1781–1868), who deduced it empirically in 1812. It is possible to prove this law rigorously from Maxwell's equations.

▶ **Example 2.** We wish to use a plate of glass ($n = 1.50$) as a polarizer. What is the polarizing angle? What is the angle of refraction?

From Eq. 46-3,

$$\theta_p = \tan^{-1} 1.50 = 56.3^\circ.$$

The angle of refraction follows from Snell's law:

$$(1) \sin \theta_p = n \sin \theta_r$$

or

$$\sin \theta_r = \frac{\sin 56.3^\circ}{1.50} = 0.555 \qquad \theta_r = 33.7^\circ.$$ ◀

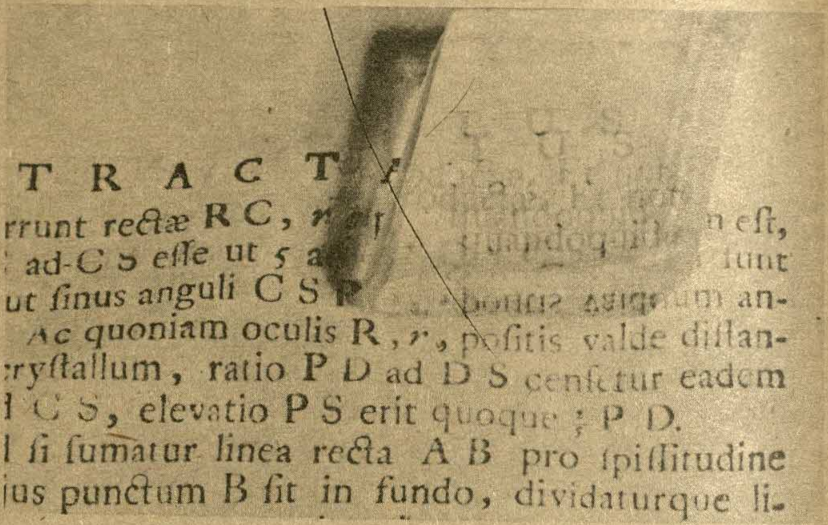
46-4 Double Refraction

In earlier chapters we assumed that the speed of light, and thus the index of refraction, is independent of the direction of propagation in the medium and of the state of polarization of the light. Liquids, amorphous solids such as glass, and crystalline solids having cubic symmetry normally show this behavior and are said to be *optically isotropic*. Many other crystalline solids are optically *anisotropic* (that is, not isotropic).*

Solids may be anisotropic in many properties. Mica cleaves readily in one plane only; a cube of crystalline graphite does not have the same electric resistance between all pairs of opposite faces; a cube of crystalline nickel magnetizes more readily in certain directions than in others, etc. If a solid is a mixture of a large number of tiny crystallites, it may appear to be isotropic because of the random orientations of the crystallites. Powdered

---

* Many transparent amorphous solids such as glasses and plastics become optically anisotropic when they are mechanically stressed. This fact is useful in engineering design studies in that strains in gears, bridge structures, etc., can be studied quantitatively by building plastic models, stressing them appropriately, and examining the optical anisotropy that results, using polarization techniques. The interested student should consult "Photoelasticity," a chapter by H. T. Jessop in Vol. 6 of the *Encyclopedia of Physics*, edited by H. Flugge (1958), Springer Verlag, Berlin.



T R A C T

rrunt rectæ R C, ... eſt,

ad-C S eſſe ut S a ... ſunt

ut ſinus anguli C S P ... an-

Ac quoniam oculis R, r, poſitis valde diſtan-

ryſtallum, ratio P D ad D S cenſetur eadem

C S, elevatio P S erit quoque P D.

ſi ſumatur linea recta A B pro ſpiſſitudine

ius punctum B ſit in fundo, dividaturque li-

Fig. 46–10 Two images, one polarized 90° relative to the other, are formed by a calcite crystal. The book on which the crystal is lying is Huygens' *Opera Reliqua*, wherein the phenomenon of birefringence is discussed.

mica, for example, compacted to a solid mass with a binder, does not exhibit the cleavage properties that characterize the crystallites making it up.

Figure 46–10, in which a polished crystal of calcite ($CaCO_3$) is laid over some printed letters, shows the optical anisotropy of this material; *the image appears double.* Figure 46–11 shows a beam of unpolarized light falling on a calcite crystal at right angles to one of its faces. The single beam splits into two at the crystal surface. The "double-bending" of a beam transmitted through calcite, exhibited in Figs. 46–10 and 46–11, is called *double refraction.*

If the two emerging beams in Fig. 46–11 are analyzed with a polarizing sheet they are found to be plane-polarized with their planes of vibration at right angles to each other, a fact discovered by Huygens in 1678. Huygens used a second calcite crystal to investigate the polarization states of the beams labeled *o* and *e* in the figure.

If experiments are carried out at various angles of incidence, one of the beams in Fig. 46–11 (represented by the *ordinary ray*, or *o*-ray) will be found to obey Snell's law of refraction at the crystal surface, just like a ray passing from one isotropic medium into another. The second beam (represented by the *extraordinary ray*, or *e*-ray) will not. In Fig. 46–11, for example, the angle of incidence for the incident light is zero but the angle of refraction of the *e*-ray, contrary to the prediction of Snell's law, is not. In general, the *e*-ray does not even lie in the plane of incidence.

This difference between the waves represented by the *o*- and *e*-rays with respect to Snell's law can be explained in these terms:

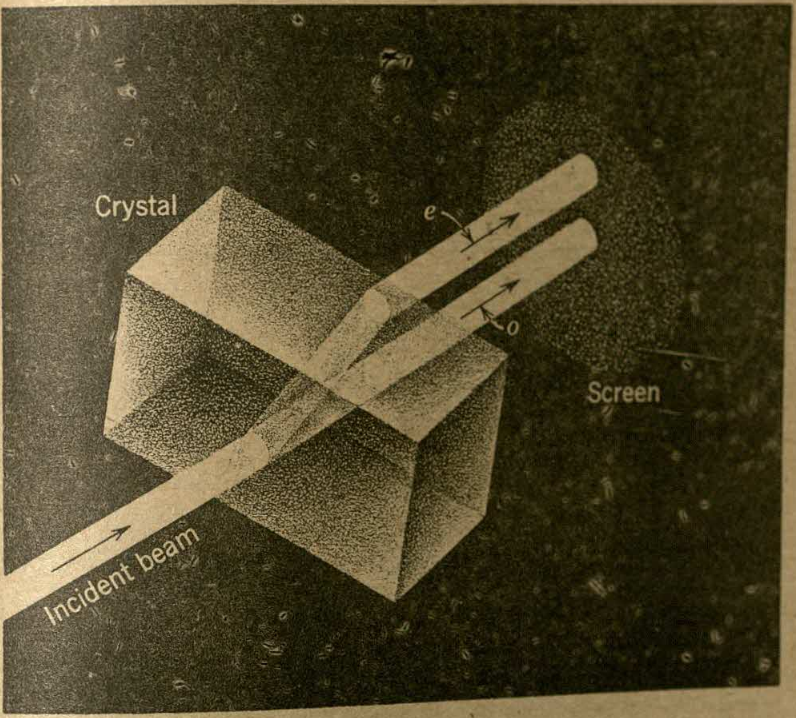


Fig. 46–11 A beam of unpolarized light falling on a calcite crystal is split into two beams which are polarized at right angles to each other.

1. The *o*-wave travels in the crystal with the same speed v_o in all directions. In other words, the crystal has, for this wave, a single index of refraction n_o, just like an isotropic solid.

2. The *e*-wave travels in the crystal with a speed that varies with direction from v_o to a larger value (for calcite) v_e. In other words, the index of refraction, defined as c/v, varies with direction from n_o to a smaller value (for calcite) n_e.

The quantities n_o and n_e are called the *principal indices of refraction* for the crystal. Problem 7 suggests how to measure them. Table 46-1 shows these indices for six doubly refracting crystals. For three of them the *e*-wave is slower; for the other three it is faster. Some doubly refracting crystals (mica, topaz, etc.) are more complex optically than calcite and require *three* principal indices of refraction for a complete description of their optical properties. Crystals whose basic crystal structure is cubic (see Fig. 45-13) are optically isotropic, requiring only *one* index of refraction.

The behavior for the speeds of the two waves traveling in calcite is summarized by Fig. 46-12, which shows two wave surfaces spreading out from an imaginary point light source S imbedded in the crystal. The *o*-wave sur-

Table 46–1

PRINCIPAL INDICES OF REFRACTION OF
SEVERAL DOUBLY REFRACTING CRYSTALS
(For sodium light, $\lambda = 5890$ A)

Crystal	Formula	n_o	n_e	$n_e - n_o$
Ice	H_2O	1.309	1.313	+0.004
Quartz	SiO_2	1.544	1.553	+0.009
Wurzite	ZnS	2.356	2.378	+0.022
Calcite	$CaCO_3$	1.658	1.486	−0.172
Dolomite	$CaO \cdot MgO \cdot 2CO_2$	1.681	1.500	−0.181
Siderite	$FeO \cdot CO_2$	1.875	1.635	−0.240

face is a sphere, as we would expect if the medium were isotropic. The *e*-wave surface is an ellipsoid of revolution about a characteristic direction in the crystal called the *optic axis*. The two wave surfaces represent light having two different polarization states. If we consider for the present only rays lying in the plane of Fig. 46–12, then (*a*) the plane of polarization for the *o*-rays is perpendicular to the figure, as suggested by the dots, and (*b*) that for the *e*-rays coincides with the plane of the figure, as suggested by the short lines. We describe the polarization states more fully at the end of this section.

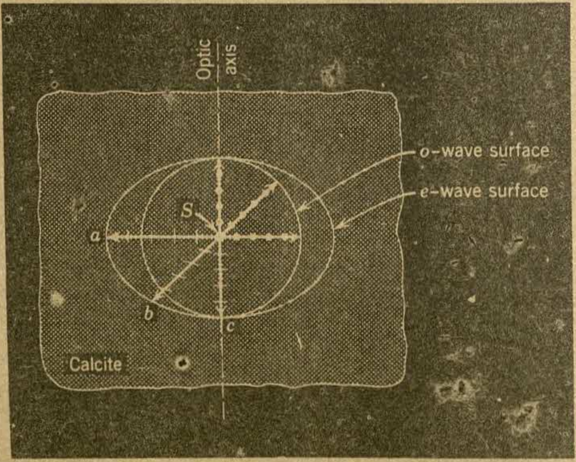


Fig. 46–12 Huygens' wave surfaces generated by a point source *S* imbedded in calcite. The polarization states for three *o*-rays and three *e*-rays are shown by the dots and lines, respectively. Note that in general (ray *Sb*) the bars representing the polarization direction are not perpendicular to the *e*-rays.

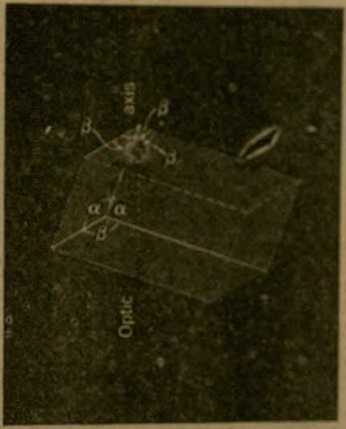


Fig. 46–13 A calcite crystal; α is 78° 13′; β is 102° 21′.

Figure 46–13, which shows a typical calcite crystal that may be obtained by cleavage from a naturally occurring crystal, shows how to locate the optic axis. The edges of calcite crystals may have any lengths but the angles at which the edges intersect always have one or another of two values, 78° 13′ or 102° 21′. The optic axis is found by erecting a line at either of the two corners where three obtuse angles meet (the "blunt" corners), making equal angles with the crystal edges. *Any line in the crystal parallel to this line is also an optic axis.*

We can use Huygens' principle to study the propagation of light waves in doubly refracting crystals. Figure 46–14*a* shows the special case in which unpolarized light falls at normal incidence on a calcite slab cut from a crystal in such a way that the optic axis is normal to the surface. Consider a wavefront that, at time $t = 0$, coincides with the crystal surface. Following Huygens, we may let any point on this surface serve as a radiating center for a double set of Huygens' wavelets, such as those in Fig. 46–12. The plane of tangency to these wavelets represents the new position of this wavefront at a later time t. The incident beam in Fig. 46–14*a* is propagated through the crystal without deviation at speed v_o. The beam emerging from the slab will have the same polarization character as the incident beam. The calcite slab, in these special circumstances only, behaves like an isotropic material, and no distinction can be made between the *o*- and the *e*-waves.

Figure 46–14*b* shows two views of another special case, namely, unpolarized incident light falling at right angles on a slab cut so that the optic axis is parallel to its surface. In this case also the incident beam is propagated without deviation. However, we can now identify *o*- and *e*-waves that travel through the crystal with different speeds, v_o and v_e, respectively. These waves are polarized at right angles to each other.

Some doubly refracting crystals have the interesting property, called *dichroism*, in which one of the polarization components is strongly absorbed within the crystal, the other being transmitted with little loss. Dichroism, illustrated in Fig. 46–15, is the basic operating principle of the commercial Polaroid sheet. The many small crystallites, imbedded in a plastic sheet with their optic axes parallel, have a polarizing action equivalent to that of a single large crystal slab.

Figure 46–14*c* shows unpolarized light falling at normal incidence on a calcite slab cut so that its optic axis makes an arbitrary angle with the crystal surface. Two spatially separated beams are produced, as in Fig. 46–11. They travel through the crystal at different speeds, that for the *o*-wave being v_o and that for the *e*-wave being intermediate between v_o and v_e.

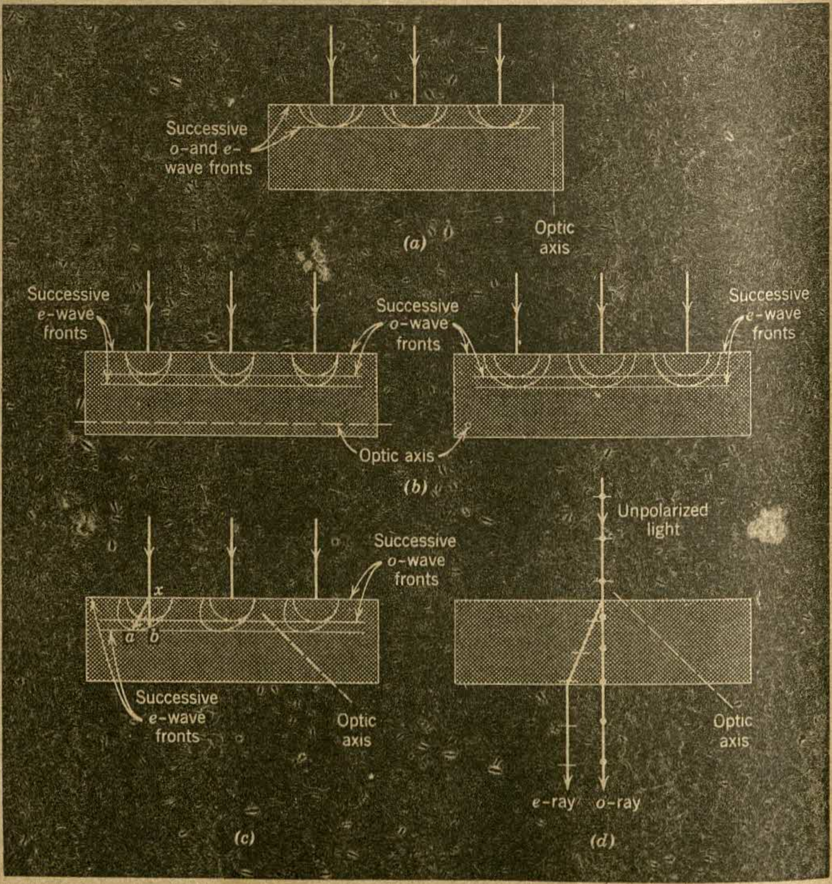


Fig. 46–14 Unpolarized light falls at normal incidence on slabs cut from a calcite crystal. The Huygens' wavelets are appropriate sections of the figure of revolution about the optic axis represented by Fig. 46–15. (*a*) No double refraction or speed difference occurs. (*b*) No double refraction occurs but there is a speed difference. (*c*) Both double refraction and a speed difference occur. (*d*) Same as (*c*) but showing the polarization states and the emerging rays.

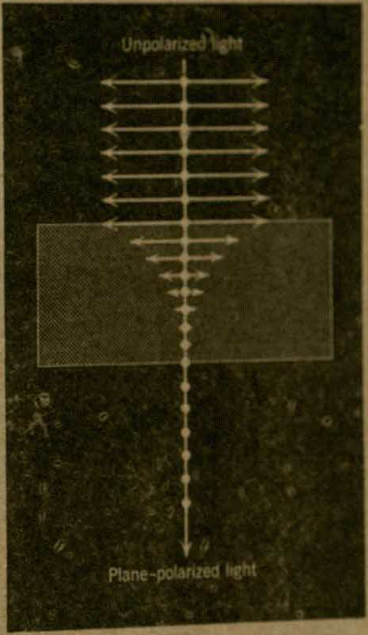


Fig. 46–15 Showing the absorption of one polarization component inside a dichroic crystal of the type used in Polaroid sheets.

Note that ray xa represents the shortest *optical* path for the transfer of light energy from point x to the *e*-wavefront. Energy transferred along any other ray, in particular along ray xb, would have a longer transit time, a consequence of the fact that the speed of *e*-waves varies with direction.* Figure 46–14*d* represents the same case as Fig. 46–14*c*. It shows the rays emerging from the slab, as in Fig. 46–11, and makes clear that the emerging beams are polarized at right angles to each other, that is, they are *cross-polarized*.

We now seek to understand, in terms of the atomic structure of optically anisotropic crystals, how cross-polarized light waves with different speeds can exist. Light is propagated through a crystal by the action of the vibrating **E** vectors of the wave on the electrons in the crystal. These electrons, which experience electrostatic restoring forces if they are moved from their equilibrium positions, are set into forced periodic oscillation about these positions and pass along the transverse wave disturbance that constitutes the light wave. The strength of the restoring forces may be measured by a force constant k, as for the simple harmonic oscillator discussed in Chapter 15 (see Eq. 15–4).

In optically isotropic materials the force constant k is the same for all directions of displacement of the electrons from their equilibrium positions. In doubly refracting crystals, however, k varies with direction. For electron displacements that lie in a plane at right angles to the optic axis k has the constant value k_o, no matter how the displacement is oriented in this plane. For displacements parallel to the

---

* The student who has not previously read Section 41–6 on Fermat's principle may care to do so now.

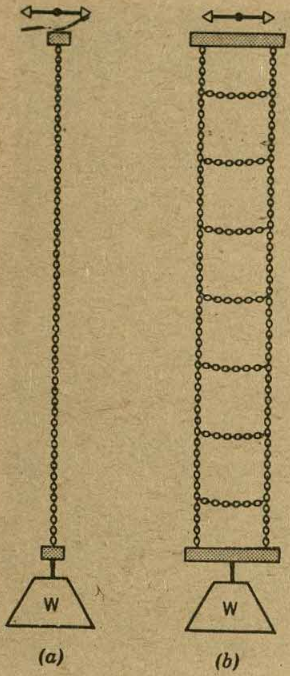


Fig. 46–16 Two views of a one-dimensional mechanical model for double refraction.

optic axis, k has the larger value (for calcite) k_e.* Note carefully that the speed of a wave in a crystal is determined by the direction in which the **E** vectors vibrate and *not* by the direction of propagation. It is the transverse **E**-vector vibrations that call the restoring forces into play and thus determine the wave speed. Note too that the stronger the restoring force, that is, the larger k, the faster the wave. For waves traveling along a stretched cord, for example, the restoring force for the transverse displacements is determined by the tension F in the cord. Equation 19–12 shows that an increase in F means an increase in the wave speed v.

Figure 46–16, a long weighted "tire chain" supported at its upper end, provides a one-dimensional mechanical analogy for double refraction. It applies specifically to *o*- and *e*-waves traveling at right angles to the optic axis, as in Fig. 46–14*b*. If the supporting block is oscillated, as in Fig. 46–16*a*, a transverse wave travels along the chain with a certain speed. If the block is oscillated lengthwise, as in Fig. 46–16*b*, another transverse wave is also propagated. The restoring force for the second wave is greater than for the first, the chain being more rigid in the plane of Fig. 46–16*b* than in the plane of Fig. 46–16*a*. Thus the second wave travels along the chain with a greater speed.

In the language of optics we would say that the speed of a transverse wave in the chain depends on the orientation of the plane of vibration of the wave. If we oscillate the top of the chain in a random way, the wave disturbance at a point along the chain can be described as the sum of two waves, polarized at right angles and traveling with different speeds. This corresponds exactly to the optical situation of Fig. 46–14*b*.

---

* For doubly refracting crystals with $n_e > n_o$ (see Table 46–1) k for displacements parallel to the optic axis is *smaller* than for those at right angles to it. Also, for crystals with three principal indices of refraction, there will be three principal force constants. Such crystals have two optic axes and are called *biaxial*. The crystals listed in Table 46–1 have only a single optic axis and are called *uniaxial*.

For waves traveling parallel to the optic axis, as in Fig. 46–14*a*, or for waves in optically isotropic materials, the appropriate mechanical analogy is a single weighted hanging chain. Here there is only one speed of propagation, no matter how the upper end is oscillated. The restoring forces are the same for all orientations of the plane of polarization of waves traveling along such a chain.

These considerations allow us to understand more clearly the polarization states of the light represented by the double-wave surface of Fig. 46–12. For the (spherical) *o*-wave surface, the **E**-vector vibrations must be everywhere at right angles to the optic axis. If this is so, the same force constant k_o will always be operative, and the *o*-waves will travel with the same speed in all directions. More specifically, if we draw a ray in Fig. 46–12 from S to the *o*-wave surface, considered three-dimensionally (that is, as a sphere), the **E**-vector vibrations will always be at right angles to the plane defined by this ray and the optic axis. Thus these vibrations will always be at right angles to the optic axis.

For the (ellipsoidal) *e*-wave surface, the **E**-vector vibrations in general have a component parallel to the optic axis. For rays such as Sa in Fig. 46–12 or for the *e*-rays of Fig. 46–14*b*, the vibrations are completely parallel to this axis. Thus a relatively strong force constant (in calcite) k_e is operative, and the wave speed v_e will be relatively high. For *e*-rays, such as Sb in Fig. 46–12, the parallel component of the **E**-vector vibrations is less than 100%, so that the corresponding wave speed will be less than v_e. For ray Sc, in Fig. 46–12, the parallel component is zero, and the distinction between *o*- and *e*-rays disappears.

46–5 Circular Polarization

Let plane-polarized light of angular frequency ω ($= 2\pi\nu$) fall at normal incidence on a slab of calcite cut so that the optic axis is parallel to the face of the slab, as in Fig. 46–17. The two waves that emerge will be plane-polarized at right angles to each other, and, if the incident plane of vibration is at 45° to the optic axis, they will have equal amplitudes. Since the waves travel through the crystal at different speeds, there will be a phase difference ϕ between them when they emerge from the crystal. If the crystal thickness is chosen so that (for a given frequency of light) $\phi = 90°$, the slab is called a *quarter-wave plate*. The emerging light is said to be *circularly polarized.*

In Section 15–7 we saw that the two emerging plane-polarized waves just described (vibrating at right angles with a 90° phase difference) can be represented as the projections on two perpendicular axes of a vector rotating with

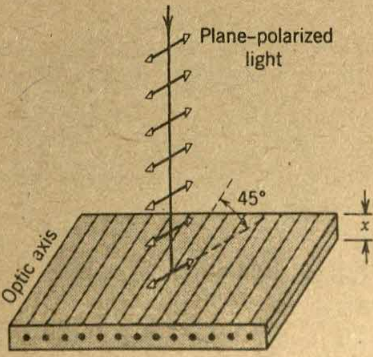


Fig. 46–17 Plane-polarized light falls on a doubly refracting slab of thickness x cut with its optic axis parallel to the surface. The plane of vibration of the incident light is oriented to make an angle of 45° with the optic axis.

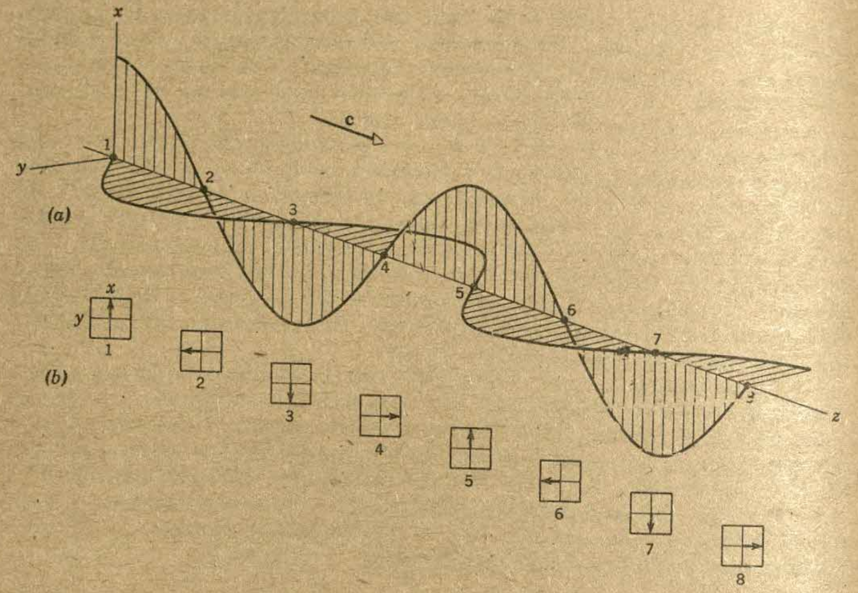


Fig. 46–18 (*a*) Two plane-polarized waves of equal amplitude and at right angles to each other are moving in the *z* direction. They differ in phase by 90°; where one wave has maximum values, the other is zero. (*b*) Views of the resultant amplitude of the approaching wave as seen by observers located at the positions shown on the *z* axis. Note that each observer will see the resultant vector rotate clockwise with time.

angular frequency ω about the propagation direction. These two descriptions of circularly polarized light are completely equivalent. Figure 46–18 clarifies the relationship between these two descriptions.

▶ **Example 3.** A quartz quarter-wave plate is to be used with sodium light (λ = 5890 A). What must its thickness be?

Two waves travel through the slab at speeds corresponding to the two principal indices of refraction given in Table 46–1 ($n_e = 1.553$ and $n_o = 1.544$). If the crystal thickness is x, the number of wavelengths of the first wave contained in the crystal is

$$N_e = \frac{x}{\lambda_e} = \frac{xn_e}{\lambda},$$

where λ_e is the wavelength of the *e*-wave in the crystal and λ is the wavelength in air. For the second wave the number of wavelengths is

$$N_o = \frac{x}{\lambda_o} = \frac{xn_o}{\lambda},$$

where λ_o is the wavelength of the *o*-wave in the crystal. The difference $N_e - N_o$ must be one-fourth, or

$$\frac{1}{4} = \frac{x}{\lambda}(n_e - n_o).$$

This equation yields

$$x = \frac{\lambda}{4(n_e - n_o)} = \frac{5890\text{ A}}{(4)(1.553 - 1.544)} = 0.016\text{ mm}.$$

This plate is rather thin; most quarter-wave plates are made from mica, splitting the sheet to the correct thickness by trial and error.

Example 4. A beam of circularly polarized light falls on a polarizing sheet. Describe the emerging beam.

The circularly polarized light, as it enters the sheet, can be represented by

$$E_x = E_m \sin \omega t$$

and

$$E_y = E_m \cos \omega t,$$

where x and y represent arbitrary perpendicular axes. These equations correctly represent the fact that a circularly polarized wave is equivalent to two plane-polarized waves with equal amplitude and a 90° phase difference.

The resultant amplitude in the incident circularly polarized wave is

$$E_{\text{cp}} = \sqrt{E_x^2 + E_y^2} = \sqrt{E_m^2 (\sin^2 \omega t + \cos^2 \omega t)} = E_m,$$

an expected result if the circularly polarized wave is represented as a rotating vector. The resultant intensity in the incident circularly polarized wave is proportional to E_m^2, or

$$I_{\text{cp}} \propto E_m^2. \tag{46-4}$$

Let the polarizing direction of the sheet make an arbitrary angle θ with the x axis as shown in Fig. 46-19. The instantaneous value of the plane-polarized wave transmitted by the sheet is

$$\begin{aligned} E &= E_y \sin \theta + E_x \cos \theta \\ &= E_m \cos \omega t \sin \theta + E_m \sin \omega t \cos \theta \\ &= E_m \sin (\omega t + \theta). \end{aligned}$$

The intensity of the wave transmitted by the sheet is proportional to E^2, or

$$I \propto E_m^2 \sin^2 (\omega t + \theta).$$

The eye and other measuring instruments respond only to the average intensity $\bar{I}$, which is found by replacing $\sin^2 (\omega t + \theta)$ by its average value over one or more cycles $(= \frac{1}{2})$, or

$$\bar{I} \propto \tfrac{1}{2} E_m^2.$$

Comparison with Eq. 46-4 shows that inserting the polarizing sheet reduces the intensity by one-half. The orientation of the sheet makes no difference, since θ does not

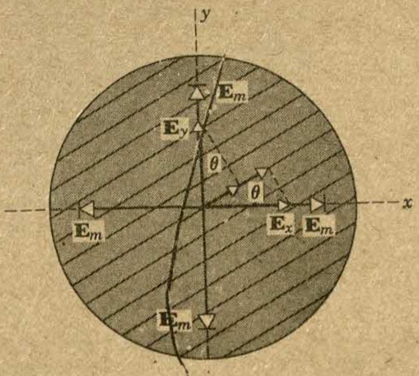


Fig. 46-19 Circularly polarized light falls on a polarizing sheet. $\mathbf{E}_x$ and $\mathbf{E}_y$ are instantaneous values of the two components, their maximum values being $\mathbf{E}_m$.

appear in this equation; this is to be expected if circularly polarized light is represented by a rotating vector, all azimuths about the propagation direction being equivalent. Inserting a polarizing sheet in an *unpolarized* beam has just the same effect, so that a simple polarizing sheet cannot be used to distinguish between unpolarized and circularly polarized light.

Example 5. A beam of light is thought to be circularly polarized. How may this be verified?

Insert a quarter-wave plate. If the beam is circularly polarized, the two components will have a phase difference of 90° between them. The quarter-wave plate will introduce a further phase difference of ±90° so that the emerging light will have a phase difference of either zero or 180°. In either case the light will now be *plane-polarized* and can be made to suffer complete extinction by rotating a polarizer in its path.

Does the quarter-wave plate have to be oriented in any particular way to carry out this test?

Example 6. A plane-polarized light wave of amplitude E_0 falls on a calcite quarter-wave plate with its plane of vibration at 45° to the optic axis of the plate, which is taken as the y axis; see Fig. 46–20. The emerging light will be circularly polarized. In what direction will the rotating electric vector appear to rotate? The direction of propagation is out of the page.

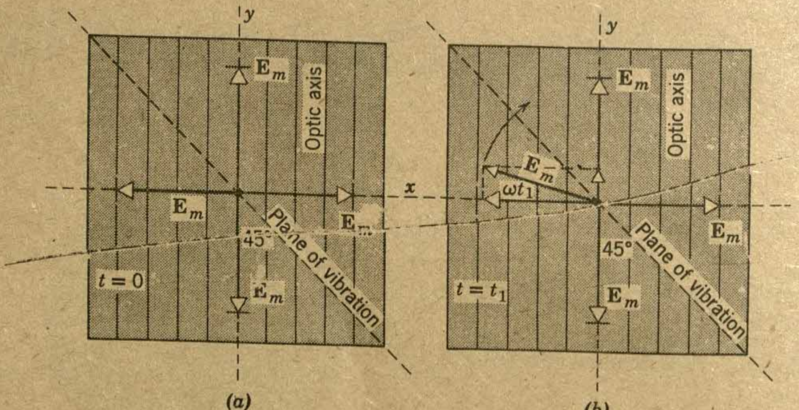
The wave component whose vibrations are parallel to the optic axis (the *e*-wave) can be represented as it emerges from the plate as

$$E_y = (E_0 \cos 45^\circ) \sin \omega t = \frac{1}{\sqrt{2}} E_0 \sin \omega t = E_m \sin \omega t.$$

The wave component whose vibrations are at right angles to the optic axis (the *o*-wave) can be represented as

$$E_x = (E_0 \sin 45^\circ) \sin (\omega t - 90^\circ) = -\frac{1}{\sqrt{2}} E_0 \cos \omega t = -E_m \cos \omega t,$$

the 90° phase shift representing the action of the quarter-wave plate. Note that E_x



(a) (b)

Fig. 46–20 Plane-polarized light falls from behind on a quarter-plate oriented so that the light emerging from the page is circularly polarized. In this case the electric vector E_m rotates clockwise as seen by an observer facing the light source.

reaches its maximum value one-fourth of a cycle *later* than E_y does, for, in calcite, wave E_x (the *o*-wave) travels *more slowly* than wave E_y (the *e*-wave).

To decide the direction of rotation, let us locate the tip of the rotating electric vector at two instants of time, (*a*) $t = 0$ and (*b*) a short time t_1 later chosen so that ωt_1 is a small angle. At $t = 0$ the coordinates of the tip of the rotating vector (see Fig. 46-20*a*) are

$$E_y = 0 \quad \text{and} \quad E_x = -E_m.$$

At $t = t_1$ these coordinates become, approximately,

$$E_y = E_m \sin \omega t_1 \cong E_m(\omega t_1)$$

$$E_x = -E_m \cos \omega t_1 \cong -E_m.$$

Figure 46-20*b* shows that the vector that represents the emerging circular polarized light is rotating clockwise; by convention such light is called *right-circularly polarized*, the observer always being considered to face the light source.

If the plane of vibration of the incident light in Fig. 46-20 is rotated through ±90°, the emerging light will be *left-circularly polarized*. ◀

46-6 Angular Momentum of Light

That light waves can deliver *linear momentum* to an absorbing screen or to a mirror is in accord with classical electromagnetism, with quantum physics, and with experiment. The facts of circular polarization suggest that light so polarized might also have *angular* momentum associated with it. This is indeed the case; once again the prediction is in accord with classical electromagnetism and with quantum physics. Experimental proof was provided in 1936 by Beth, who showed that when circularly polarized light is produced in a doubly refracting slab the slab experiences a reaction torque.

The angular momentum carried by light plays a vital role in understanding the emission of light from atoms and of γ-rays from nuclei. If light carries away angular momentum as it leaves the atom, the angular momentum of the residual atom must change by exactly the amount carried away; otherwise the angular momentum of the isolated system *atom plus light* will not be conserved.

Classical and quantum theory both predict that if a beam of circularly polarized light is completely absorbed by an object on which it falls, an angular momentum given by

$$L = \frac{U}{\omega} \tag{46-5}$$

is transferred to the object, where U is the amount of absorbed energy and ω the angular frequency of the light. The student should verify that the dimensions in Eq. 46-5 are consistent.

46-7 Scattering of Light

A light wave, falling on a transparent solid, causes the electrons in the solid to oscillate periodically in response to the time-varying electric vector of the incident wave. The wave that travels through the medium is the resultant of the incident wave and of the radiations from the oscillating electrons. The resultant wave has a maximum intensity in the direction of the incident beam, falling off rapidly on either side. The lack of sideways scattering, which would be essentially complete in a large "perfect" crystal, comes about because the oscillating charges in the medium act cooperatively or coherently.

When light passes through a gas, we find much more sideways scattering. The oscillating electrons in this case, being separated by relatively large distances and not being bound together in a rigid structure, act independently rather than cooperatively. Thus the rigid cancellation of wave disturbances that are not in the forward direction is less likely to occur; there is more sideways scattering.

Light scattered sideways from a gas can be wholly or partially polarized, even though the incident light is unpolarized. Figure 46–21 shows an unpolarized beam moving upward on the page and striking a gas atom at *a*. The electrons at *a* will oscillate in response to the electric components of the incident wave, their motion being equivalent to two oscillating dipoles whose axes are represented by the arrow and the dot at *a*. An oscillating dipole does not radiate along its own line of action. Thus an observer at *b* would receive no radiation from the dipole represented by the arrow at *a*. The radiation reaching him would come entirely from the dipole represented by the dot at *a*; thus this radiation would be plane-polarized, the plane of vibration passing through the line *ab* and being normal to the page.

Observers at *c* and *d* would detect partially polarized light, since the dipole represented by the arrow at *a* would radiate somewhat in these directions. Observers viewing the transmitted or the back-scattered light would not detect any polarization effects because both dipoles at *a* would radiate equally in these two directions.

A familiar example is the scattering of sunlight by the molecules of the earth's atmosphere. If the atmosphere were not present, the sky would appear black except when we looked directly at the sun This has been verified

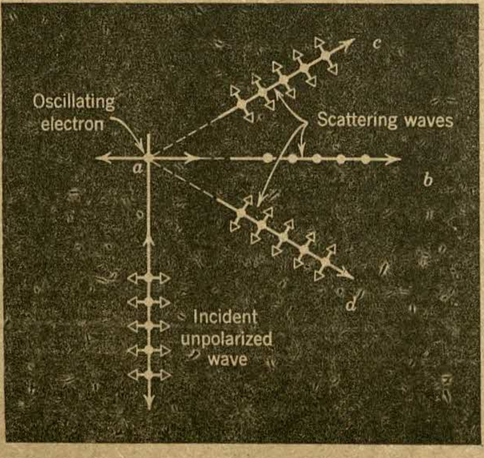


Fig. 46–21 Light is polarized either partially (*c* and *d*) or completely (*b*) by scattering from a gas molecule at *a*.

by measurements made in rockets and satellites above the atmosphere. We can easily check with a polarizer that the light from the cloudless sky is at least partially polarized. This fact is used in polar exploration in the so-called *solar compass*. In this device we establish direction by noting the nature of the polarization of the scattered sunlight. As is well known, magnetic compasses are not useful in these regions. It has been learned * that bees orient themselves in their flights between their hive and the pollen sources by means of polarization of the light from the sky; bees' eyes contain built-in polarization-sensing devices.

It still remains to be explained why the light scattered from the sky is predominantly blue and why the light received directly from the sun—particularly at sunset when the length of the atmosphere that it must traverse is greatest—is red. The cross section of an atom or molecule for light scattering depends on the wavelength, blue light being scattered more effectively than red light. Since the blue light is largely scattered, the transmitted light will have the color of normal sunlight with the blues largely removed; it is therefore more reddish in appearance.

The fact that the scattering cross section for blue light is higher than that for red light can be made reasonable. An electron in an atom or molecule is bound there by strong restoring forces. It has a definite natural frequency, like a small mass suspended in space by an assembly of springs. The natural frequency for electrons in atoms and molecules is usually in a region corresponding to violet or ultraviolet light.

When light is allowed to fall on such bound electrons, it sets up forced oscillations at the frequency of the incident light beam. In mechanical resonant systems it is possible to "drive" the system most effectively if we impress on it an external force whose frequency is as close as possible to that of the natural resonant frequency. In the case of light the blue is closer to the natural resonant frequency of the bound electron than is the red light. Therefore, we would expect the blue light to be more effective in causing the electron to oscillate, and thus it will be more effectively scattered.

46-8 Double Scattering

When X-rays were discovered in 1898, there was much speculation whether they were waves or particles. In 1906 they were established as transverse waves by Charles Glover Barkla (1877–1944) by means of a polarization experiment.

When the unpolarized X-rays strike scattering block S_1 in Fig. 46–22, they set the electrons into oscillatory motion. The considerations of the preceding section require that the X-rays scattered toward the second block be plane-polarized as shown in the figure. Let this wave be scattered from the second scattering block, and let us examine the radiation scattered from it by rotating a detector D in a plane at right angles to the line joining the blocks. The electrons will oscillate parallel to each other, and the positions of maximum and zero intensity will be as shown. A plot of detector reading as a function of the angle ϕ supports the hypothesis that X-rays are transverse waves. If the X-rays were a stream of particles or a longitudinal wave, these effects could by no means be so readily understood. Thus Barkla's important experiment established that X-rays are a part of the electromagnetic spectrum.

---

* See *Scientific American*, July 1955, and *Bees: Their Vision, Chemical Sense, and Language*, K. von Frisch, Cornell University Press, 1950.

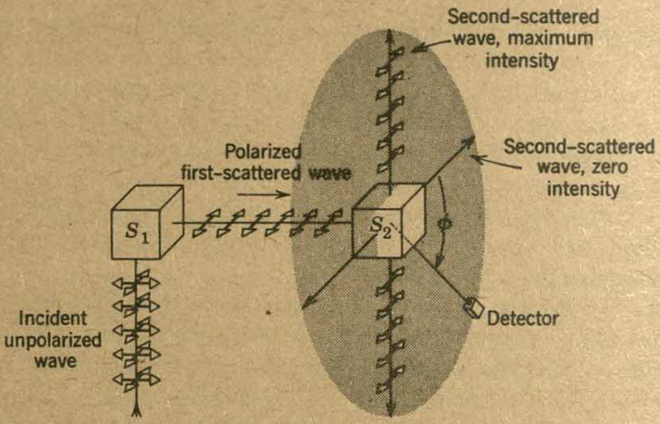

Second-scattered wave, maximum intensity
Second-scattered wave, zero intensity
Polarized first-scattered wave
S_1
S_2
ϕ
Incident unpolarized wave
Detector

Fig. 46–22 A double-scattering experiment used by Barkla to show that X-rays are transverse waves.

In his later studies the student will learn that beams of particles such as electrons, protons, and pions can be viewed as waves. Scattering (including double scattering) techniques are often used to investigate the polarization characteristics of such beams.

QUESTIONS

1. Why do sunglasses made of polarizing materials have a marked advantage over those that simply depend on absorption effects?

2. Unpolarized light falls on two polarizing sheets so oriented that no light is transmitted. If a third polarizing sheet is placed between them, can light be transmitted?

3. Can polarization by reflection occur if the light is incident on the interface from the side with the higher index of refraction (glass to air, for example)?

4. Is the optic axis of a doubly refracting crystal simply a line or a direction in space? Has it a direction sense, like an arrow? What about the characteristic direction of a polarizing sheet?

5. Devise a way to identify the polarizing direction of a sheet of Polaroid.

6. If ice is doubly refracting (see Table 46–1), why don't we see two images of objects viewed through an ice cube?

7. From Table 46–1, would you expect a quarter-wave plate made from calcite to be thicker than one made from quartz?

8. Does the *e*-wave in doubly refracting crystals always travel at a speed given by c/n_e?

9. In Fig. 46–14*a* and *b* describe qualitatively what happens if the incident beam falls on the crystal with an angle of incidence that is not zero. Assume in each case that the incident beam remains in the plane of the figure.

10. Devise a way to identify the direction of the optic axis in a quarter-wave plate.

11. If plane-polarized light falls on a quarter-wave plate with its plane of vibration making an angle of (*a*) 0° or (*b*) 90° with the axis of the plate, describe the transmitted light. (*c*) If this angle is arbitrarily chosen, the transmitted light is called *elliptically polarized;* describe such light.

12. What would be the action of a *half-wave plate* (that is, a plate twice as thick as a

quarter-wave plate) on (*a*) plane-polarized light (assume the plane of vibration to be at 45° to the optic axis of the plate), (*b*) circularly polarized light, and (*c*) unpolarized light?

13. You are given an object which may be (*a*) a disk of grey glass, (*b*) a polarizing sheet, (*c*) a quarter-wave plate, or (*d*) a half-wave plate (see Question 12). How could you identify it?

14. Can a plane-polarized light beam be represented as a sum of two circularly polarized light beams of opposite rotation? What effect has changing the phase of one of the circular components on the resultant beam?

15. How can a right-circularly polarized light beam be transformed into a left-circularly polarized beam?

16. Could (*a*) a radar beam and (*b*) a sound wave in air be circularly polarized?

17. A beam of light is said to be unpolarized, plane-polarized, or circularly polarized. How could you choose among them experimentally?

18. A parallel beam of light is absorbed by an object placed in its path. Under what circumstances will (*a*) linear momentum and (*b*) angular momentum be transferred to the object?

19. When observing a clear sky through a polarizing sheet, one finds that the intensity varies by a factor of two on rotating the sheet. This does not happen when one views a cloud through the sheet. Can you devise an explanation?

PROBLEMS

1. Unpolarized light falls on two polarizing sheets placed one on top of the other. What must be the angle between the characteristic directions of the sheets if the intensity of the transmitted light is (*a*) one-third the maximum intensity of the transmitted beam or (*b*) one-third the intensity of the incident beam? Assume that the polarizing sheet is ideal, that is, that it reduces the intensity of unpolarized light by exactly 50%.

2. An unpolarized beam of light is incident on a group of four polarizing sheets which are lined up so that the characteristic direction of each is rotated by 30° clockwise with respect to the preceding sheet. What fraction of the incident intensity is transmitted?

3. Describe the state of polarization represented by these sets of equations:

$$(a)\quad E_x = E\sin(kz - \omega t)$$

$$E_y = E\cos(kz - \omega t),$$

$$(b)\quad E_x = E\cos(kz - \omega t)$$

$$E_y = E\cos\left(kz - \omega t + \frac{\pi}{4}\right),$$

$$(c)\quad E_x = E\sin(kz - \omega t)$$

$$E_y = -E\sin(kz - \omega t).$$

4. (*a*) At what angle of incidence will the light reflected from water be completely polarized? (*b*) Does this angle depend on the wavelength of the light?

5. Calculate the range of polarizing angles for white light incident on fused quartz. Assume that the wavelength limits are 4000 and 7000 A and use the dispersion curve of Fig. 41–2.

6. A narrow beam of unpolarized light falls on a calcite crystal cut with its optic axis as shown in Fig. 46–23. (*a*) For $t = 1.0$ cm and for $\theta_i = 45°$, calculate the perpendicular distance between the two emerging rays *x* and *y*. (*b*) Which is the *o*-ray and which the *e*-ray? (*c*) What are the states of polarization of the emerging rays? (*d*) Describe what

happens if a polarizer is placed in the incident beam and rotated. (Hint: Inside the crystal the **E**-vector vibrations for one ray are always perpendicular to the optic axis and for the other ray they are always parallel. The two rays are described by the indices n_o and n_e; *in this plane* each ray obeys Snell's law.)

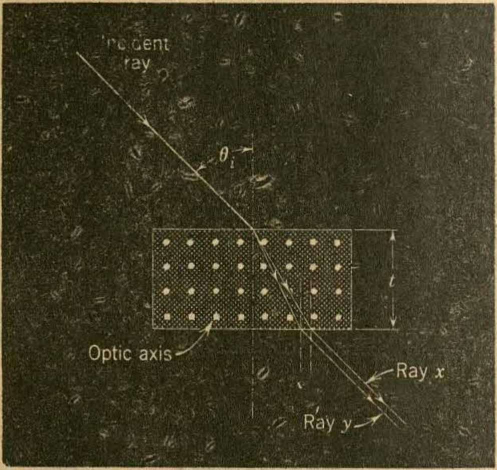


Fig. 46–23

7. A prism is cut from calcite so that the optic axis is parallel to the prism edge as shown in Fig. 46–24. Describe how such a prism might be used to measure the two principal indices of refraction for calcite. (Hint: See hint in Problem 6; see also Example 3, Chapter 41.)

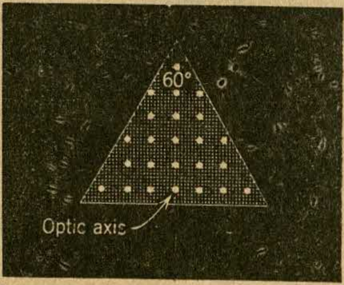


Fig. 46–24

8. How thick must a sheet of mica be if it is to form a quarter-wave plate for yellow light ($\lambda = 5890$ A)? Mica cleaves in such a way that the appropriate indices of refraction, for transmission at right angles to the cleavage plane, are 1.6049 and 1.6117.

9. Prove that two plane-polarized light waves of equal amplitude, their planes of vibration being at right angles to each other, cannot produce interference effects. (Hint: Prove that the intensity of the resultant light wave, averaged over one or more cycles of oscillation, is the same no matter what phase difference exists between the two waves.)

10. Show that in a parallel beam of circularly polarized light the angular momentum per unit volume L_v is given by

$$L_v = \frac{P}{\omega c},$$

where P is the power per unit area (watts/cm^2, say) of the beam. Start from Eq. 46–5.

11. Assume that a parallel beam of circularly polarized light whose intensity is 100 watts is absorbed by an object. At what rate is angular momentum transferred to the object? If the object is a flat disk of diameter 5.0 mm and mass 1.0×10^{-2} gm, after how long a time (assuming it is free to rotate about its axis) would it attain an angular speed of 1.0 rev/sec? Assume a wavelength of 5000 A.

Light and Quantum Physics

CHAPTER 47

47–1 Sources of Light

We have studied the propagation, reflection, refraction, diffraction, polarization, scattering, and interference of light. This chapter deals in part with the *production* of light and with the way that such studies led, in 1900, to the birth of modern quantum physics.

The most common light sources are heated solids and gases through which an electric discharge is passing. The tungsten filament of an incandescent lamp and the familiar neon sign are examples in each category. By analyzing the light from a source with a spectrometer, we can learn how strongly it radiates at various wavelengths. Figure 47–1, which is typical of spectra for heated solids, shows the results of such measurements for a heated tungsten ribbon at 2000°K.

The ordinate $\mathcal{R}_\lambda$ in Fig. 47–1 is called the *spectral radiancy,* defined so that the quantity $\mathcal{R}_\lambda\, d\lambda$ is the rate at which energy is radiated per unit area of surface for wavelengths lying in the interval λ to $\lambda + d\lambda$. Typical units for $\mathcal{R}_\lambda$ are watts/cm^2-μ; the corresponding units of $\mathcal{R}_\lambda\, d\lambda$ are watts/cm^2. In measuring $\mathcal{R}_\lambda$, all radiation emerging into the forward hemisphere is included.

Sometimes we wish to discuss the radiated energy without regard to its wavelength. An appropriate quantity here is the *radiancy* $\mathcal{R}$, defined as the rate per unit surface area at which energy is radiated into the forward hemisphere, appropriate units being watts/cm^2. It can be found by integrating the radiation present in all wavelength intervals:

$$\mathcal{R} = \int_0^\infty \mathcal{R}_\lambda\, d\lambda.$$

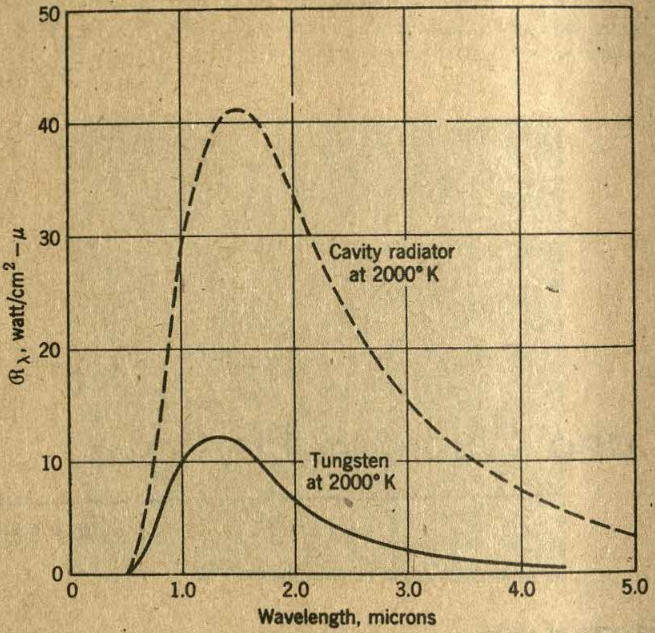


Fig. 47–1 The spectral radiancy of tungsten at 2000°K. The dashed curve refers to a cavity radiator at the same temperature. One micron (1 μ) = 10^{-6} meter = 10^4 A.

The radiancy $\mathcal{R}$ can be interpreted as the area under the plot of $\mathcal{R}_\lambda$ against λ. In Fig. 47–1 this area—and thus $\mathcal{R}$—is 23.5 watts/cm^2. The student should note the formal similarity between such curves and the Maxwell speed distribution curve of Section 24–2.

For every material there exists a family of spectral radiancy curves like that of Fig. 47–1, one curve for every temperature. If such families of curves are compared, no obvious regularities stand out. A quantitative understanding in terms of a basic theory presents serious difficulties. Fortunately, it is possible to work with an *idealized heated solid*, called a *cavity radiator*. Its light-emitting properties prove to be independent of any particular material and to vary in a simple way with temperature. In much the same way it proved convenient earlier to deal with an ideal gas rather than to analyze the properties of the infinite variety of real gases. The cavity radiator is the *ideal solid* as far as its light-emitting characteristics are concerned. We shall describe in the next two sections how the theoretical study of cavity radiation in 1900 by the German physicist Max Planck (1858–1947) laid the foundations of modern quantum physics.

47–2 Cavity Radiators

Let us construct a cavity in each of three metal blocks through the walls of which a small hole is drilled. Let the blocks be made of any suitable materials; for example, tungsten, tantalum, and molybdenum. Let each block be raised to the same uniform temperature (say 2000°K) as determined by a suitable thermometer. Finally, let us observe the blocks by their emitted light in a dark room. Measurements of $\mathcal{R}$ and $\mathcal{R}_\lambda$ show the following:

1. The radiation from the cavity interior is always more intense than the radiation from the outside wall. Comparison of the two curves in Fig. 47-1 makes this clear for tungsten. For the three materials given, at 2000°K the ratio of the radiancy for the outside surface to that for the cavity is 0.259 (tungsten), 0.212 (molybdenum), and 0.232 (tantalum).

2. At a given temperature the radiancy of the hole is *identical for all three radiators,* in spite of the fact that the radiancies of the outer surfaces are different. At 2000°K the cavity radiancy (that is, the hole radiancy) is 90.0 watts/cm^2.

3. In contrast to the radiancy of the outer surfaces, the cavity radiancy $\mathcal{R}_c$ varies with temperature in a simple way, namely as

$$\mathcal{R}_c = \sigma T^4, \tag{47-2}$$

where σ is a universal constant (the Stefan-Boltzmann constant) whose measured value is 5.67×10^{-8} watt/(meter^2)(°K^4). The radiancy of the outer surfaces varies with temperature in a more complicated way and is different for different materials. It is often written as

$$\mathcal{R} = e\mathcal{R}_c = e\sigma T^4, \tag{47-3}$$

where e, the *emissivity,* depends on the material and the temperature.

4. $\mathcal{R}_\lambda$ for the cavity radiation varies with temperature in the way shown in Fig. 47-2. These curves depend only on the temperature and are quite independent of the material and of the shape and size of the cavity.

Figure 47-3 shows an actual cavity, consisting of a hollow thin-walled cylinder of tungsten heated by sending an electric current through it. The cylinder is mounted in an evacuated glass bulb, and a tiny hole is drilled

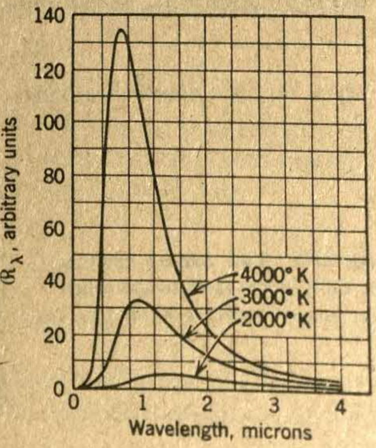


Fig. 47-2 The spectral radiancy for cavity radiation at three different temperatures.

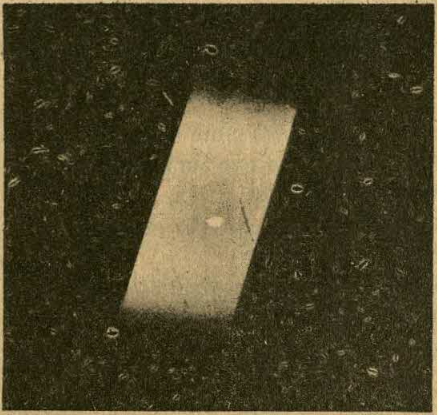


Fig. 47–3 Photograph of an incandescent tungsten tube with a small hole drilled in its wall. The radiation emerging from the hole is cavity radiation.

through the cylinder wall. It is clear from the photograph that the radiancy of the cavity interior is greater than that of the cavity walls.

Many of the facts just given about cavity radiation can be deduced from Fig. 47–4, which shows two cavities made of different materials, of arbitrary shapes, and with the same wall temperature T. Radiation, described by $\mathcal{R}_A$, goes from cavity A to cavity B and radiation described by $\mathcal{R}_B$ moves in the opposite direction. If these two rates of energy transfer are not equal, one end of the composite block will start to heat up and the other end will start to cool down, which is a violation of the second law of thermodynamics. (Why?) Thus we must have

$$\mathcal{R}_A = \mathcal{R}_B = \mathcal{R}_c, \tag{47–4}$$

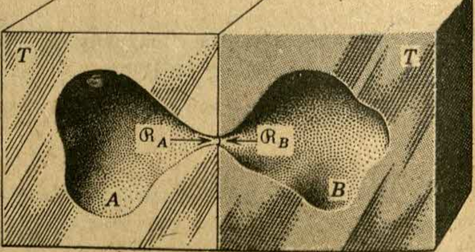
where $\mathcal{R}_c$ describes the total radiation for *all* cavities.

Not only the total radiation but also the distribution of radiant energy with wavelength must be the same for each cavity in Fig. 47–4. This can be shown by placing a filter between the two cavity openings, so chosen that it permits only a selected narrow band of wavelengths to pass. Applying the same argument, we can show that we must have

$$\mathcal{R}_{\lambda A} = \mathcal{R}_{\lambda B} = \mathcal{R}_{\lambda c}, \tag{47–5}$$

where $\mathcal{R}_{\lambda c}$ is a spectral radiancy characteristic of all cavities.

Fig. 47–4 Two radiant cavities initially at the same temperature are placed together as shown.



47-3 Planck's Radiation Formula

A theoretical explanation for the cavity radiation was the outstanding unsolved problem in physics during the years before the turn of the present century. A number of capable physicists advanced theories based on classical physics, which, however, had only limited success. Figure 47-5, for example, shows the theory of Wien; the fit to the experimental points is reasonably good, within the experimental error of the data, but definitely not exact. Wien's formula is

$$\mathcal{R}_\lambda = \frac{c_1}{\lambda^5} \frac{1}{e^{c_2/\lambda T}},$$

where c_1 and c_2 are constants that must be determined empirically by fitting the theoretical formula to the experimental data.

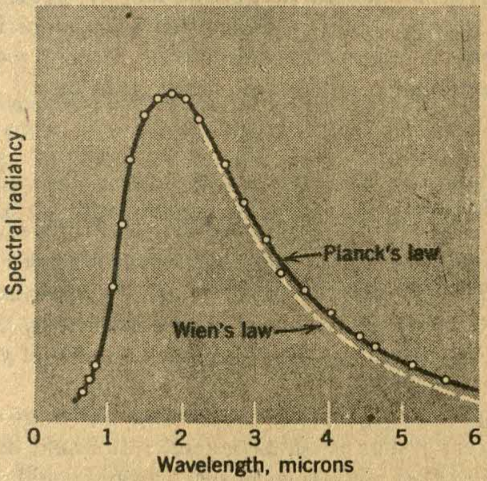


Fig. 47-5 The circles show the experimental spectral radiancy data of Coblentz for cavity radiation. The theoretical formulas of Wien and Planck are also shown, Planck's providing an excellent fit to the data.

In 1900 Max Planck pointed out that if Wien's formula were modified in a simple way it would prove to fit the data precisely. Planck's formula, announced to the Berlin Physical Society on October 19, 1900, was

$$\mathcal{R}_\lambda = \frac{c_1}{\lambda^5} \frac{1}{e^{c_2/\lambda T} - 1}. \tag{47-6}$$

This formula, though interesting and important, was still empirical at that stage and did not constitute a theory.

Planck sought such a theory in terms of a detailed model of the atomic processes taking place at the cavity walls. He assumed that the atoms that make up these walls behave like tiny electromagnetic oscillators, each with a characteristic frequency of oscillation. The oscillators emit electromagnetic energy into the cavity and absorb electromagnetic energy from it. Thus it should be possible to deduce the characteristics of the cavity radiation from those of the oscillators with which it is in equilibrium.

Planck was led to make two radical assumptions about the atomic oscillators. As eventually formulated, these assumptions are the following:

1. An oscillator cannot have *any* energy but only energies given by *

$$E = nh\nu, \tag{47-7}$$

where ν is the oscillator frequency, h is a constant (now called *Planck's constant*), and n is a number (now called a *quantum number*) that can take on only integral values. Equation 47–7 asserts that the oscillator energy is *quantized.*

2. The oscillators do not radiate energy continuously, but only in "jumps," or *quanta.* These quanta of energy are emitted when an oscillator changes from one to another of its quantized energy states. Thus, if n changes by one unit, Eq. 47–7 shows that an amount of energy given by

$$\Delta E = \Delta n h\nu = h\nu \tag{47-8}$$

is radiated. As long as an oscillator remains in one of its quantized states (or *stationary states* as they are called), it neither emits nor absorbs energy.

These assumptions were radical ones and, indeed, Planck himself resisted accepting them wholeheartedly for many years. In his words, "My futile attempts to fit the elementary quantum of action [that is, the quantity h] somehow into the classical theory continued for a number of years, and they cost me a great deal of effort."

Consider the application of Planck's hypotheses to a large-scale oscillator such as a mass-spring system or an LC circuit. It would be a stoutly defended common belief that oscillations in such systems could take place with *any* value of total energy and not with only certain discrete values. In the decay of such oscillations (by friction in the mass-spring system or by resistance and radiation in the LC circuit), it would seem that the mechanical or electromagnetic energy would decrease in a perfectly continuous way and not by "jumps." There is no basis in everyday experience, however, to dismiss Planck's assumptions as violations of "common sense," for Planck's constant proves to have a very small value, namely

$$h = 6.625 \times 10^{-34} \text{ joule-sec.}$$

The following example makes this clear.

---

* Later developments show that the correct formula for a harmonic oscillator is

$$E = (n + \tfrac{1}{2})h\nu.$$

This change makes no difference to Planck's conclusions, however.

▶ **Example 1.** A mass-spring system has a mass $m = 1.0$ kg and a spring constant $k = 20$ nt/meter and is oscillating with an amplitude of 1.0 cm. (*a*) If its energy is quantized according to Eq. 47-7, what is the quantum number n? (*b*) If n changes by unity, what fractional change in energy occurs?

(*a*) From Eq. 15-8 the frequency is

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{m}} = \frac{1}{2\pi}\sqrt{\frac{20\ \text{nt/meter}}{1.0\ \text{kg}}} = 0.71\ \text{cycles/sec.}$$

From Eq. 8-8 the mechanical energy is

$$E = \tfrac{1}{2}kx_{\max}^2 = \tfrac{1}{2}(20\ \text{nt/meter})(10^{-2}\ \text{meter})^2 = 1.0 \times 10^{-3}\ \text{joule.}$$

From Eq. 47-7 the quantum number is

$$n = \frac{E}{h\nu} = \frac{1.0 \times 10^{-3}\ \text{joule}}{(6.6 \times 10^{-34}\ \text{joule-sec})(0.71\ \text{cycles/sec})} = 2.1 \times 10^{30}.$$

(*b*) If n changes by unity, the fractional change in energy is given by dividing Eq. 47-8 by Eq. 47-7, or

$$\frac{\Delta E}{E} = \frac{h\nu}{nh\nu} = \frac{1}{n} = \sim 10^{-30}.$$

Thus for large-scale oscillators the quantum numbers are enormous and the quantized nature of the energy of the oscillations will not be apparent. Similarly, we are not aware in large-scale experiments of the discrete nature of mass and the quantized nature of charge, that is, of the existence of atoms and electrons. ◀

On the basis of his two assumptions, Planck was able to derive his radiation law (Eq. 47-6) entirely from theory, receiving a Nobel prize for this accomplishment in 1918. His theoretical expressions for the hitherto empirical constants c_1 and c_2 were

$$c_1 = 2\pi c^2 h \qquad \text{and} \qquad c_2 = \frac{hc}{k},$$

where k is Boltzmann's constant (see Section 23-5) and c is the speed of light. By inserting the experimental values for c_1 and c_2, Planck was able to derive the values of both h and k. Planck described his theory to the Berlin Physical Society on December 14, 1900. Quantum physics dates from that day. Planck's ideas soon received re-enforcement from Einstein, who, in 1905, applied the concepts of energy quantization to a new area of physics, the photoelectric effect.

Before discussing this effect, it is important to realize that although Planck had quantized the energies of the oscillators in the cavity walls he *still treated the radiation within the cavity* as an electromagnetic wave. Einstein's analysis of the photoelectric effect first pointed out the inadequacy of the wave picture of light in certain situations.

47-4 Photoelectric Effect

Figure 47-6 shows an apparatus used to study the photoelectric effect. Monochromatic light, falling on metal plate A, will liberate *photoelectrons*, which can be detected as a current if they are attracted to metal cup B by

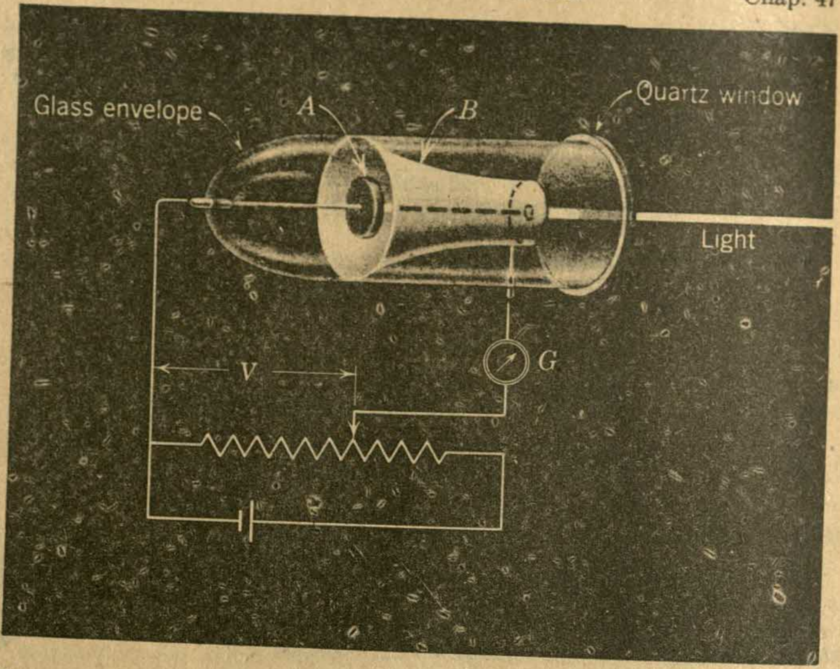


Fig. 47–6 An apparatus used to study the photoelectric effect. V cannot only be varied continuously but can be reversed in sign by a switching arrangement not shown.

means of a potential difference V applied between A and B. Galvanometer G serves to measure this *photoelectric current.*

Figure 47–7 (curve a) is a plot of the photoelectric current in an apparatus like that of Fig. 47–6, as a function of the potential difference V. If V is made large enough, the photoelectric current reaches a certain limiting value at which *all* photoelectrons ejected from plate A are collected by cup B.

If V is reversed in sign, the photoelectric current does not immediately drop to zero, which proves that the electrons are emitted from A with a finite velocity. Some will reach cup B in spite of the fact that the electric field opposes their motion. However, if this reversed potential difference is made large enough, a value V_0 (the *stopping potential*) is reached at which the photoelectric current does drop to zero. This potential difference V_0, multiplied by electron charge, measures the kinetic energy $K_{\max}$ of the *fastest* ejected photoelectron. In other words,

$$K_{\max} = eV_0. \tag{47–9}$$

Here $K_{\max}$ turns out to be independent of the intensity of the light as shown by curve b in Fig. 47–7, in which the light intensity has been reduced to one-half.

Figure 47–8 shows the stopping potential V_0 as a function of the frequency of the incident light for sodium. Note that there is a definite cutoff frequency

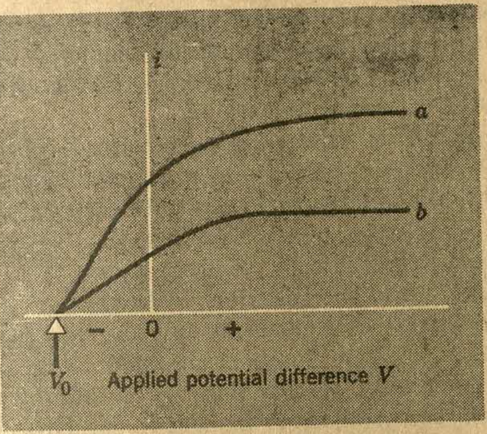


Fig. 47-7 Some data taken with the apparatus of Fig. 47-6. The applied potential difference V is called positive when the cup B in Fig. 47-6 is positive with respect to the photoelectric surface A. In curve b the incident light intensity has been reduced to one-half that of curve a.

ν_0, below which no photoelectric effect occurs. These data were taken by R. A. Millikan (1868–1953), whose painstaking work on the photoelectric effect won him the Nobel prize in 1923. Because the photoelectric effect is largely a surface phenomenon, it is necessary to avoid oxide films, grease, or other surface contaminants. Millikan devised a technique to cut shavings from the metal surface under vacuum conditions, a "machine shop *in vacuo*" as he called it.

Three major features of the photoelectric effect cannot be explained in terms of the wave theory of light:

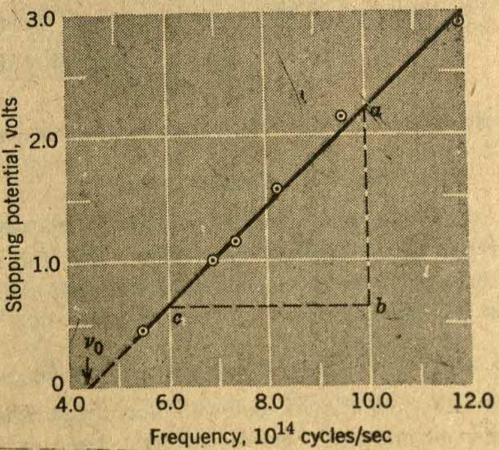


Fig. 47-8 A plot of Millikan's measurements of the stopping potential at various frequencies for sodium. The cutoff frequency ν_0 is 4.39×10^{14} cycles/sec.

1. Wave theory suggests that the kinetic energy of the photoelectrons should increase as the light beam is made more intense. However, Fig. 47–7 shows that $K_{\max}$ ($= eV_0$) *is independent of the light intensity*; this has been tested over a range of intensities of 10^7.

2. According to the wave theory, the photoelectric effect should occur for *any* frequency of the light, provided only that the light is intense enough. However, Fig. 47–8 shows that there exists, for each surface, *a characteristic cutoff frequency* ν_0. For frequencies less than this, the photoelectric effect disappears, no matter how intense the illumination.

3. If the energy of the photoelectrons is "soaked up" from the incident wave by the metal plate, it is not likely that the "effective target area" for an electron in the metal is much more than a few atomic diameters. Thus, if the light is feeble enough, there should be a measurable time lag (see Example 2) between the impinging of the light on the surface and the ejection of the photoelectron. During this interval the electron should be "soaking up" energy from the beam until it had accumulated enough energy to escape. However, *no detectable time lag has ever been measured*. This disagreement is particularly striking when the photoelectric substance is a gas; under these circumstances the energy of the emitted photoelectron must certainly be "soaked out of the beam" by a single atom.

► **Example 2.** A metal plate is placed 5 meters from a monochromatic light source whose power output is 10^{-3} watt. Consider that a given ejected photoelectron may collect its energy from a circular area of the plate as large as ten atomic diameters (10^{-9} meter) in radius. The energy required to remove an electron through the metal surface is about 5.0 ev. *Assuming light to be a wave*, how long would it take for such a "target" to soak up this much energy from such a light source?

The target area is π (10^{-9} meter)^2 or 3×10^{-18} meter^2; the area of a 5-meter sphere centered on the light source is 4π (5 meters)$^2 \cong 300$ meters^2. Thus, if the light source radiates uniformly in all directions, the rate P at which energy falls on the target is given by

$$P = (10^{-3}\ \text{watt}) \left(\frac{3 \times 10^{-18}\ \text{meter}^2}{300\ \text{meters}^2}\right) = 10^{-23}\ \text{joule/sec}.$$

Assuming that all this power is absorbed, we may calculate the time required from

$$t = \left(\frac{5\ \text{ev}}{10^{-23}\ \text{joule/sec}}\right)\left(\frac{1.6 \times 10^{-19}\ \text{joule}}{1\ \text{ev}}\right) \cong 20\ \text{hr}.$$

However, *no detectable time lag can be measured under any circumstances.* ◄

47–5 Einstein's Photon Theory

Einstein succeeded in explaining the photoelectric effect by making a remarkable assumption, namely, that the energy in a light beam travels through space in concentrated bundles, called *photons*. The energy E of a single photon (see Eq. 47–8) is given by

$$E = h\nu. \qquad (47\text{–}10)$$

Recall that Planck believed that light, although emitted from its source discontinuously, travels through space as an electromagnetic wave. Einstein's hypothesis suggests that light traveling through space behaves not like a wave at all but like a particle. Millikan, whose experiments verified Einstein's ideas in every detail, spoke of Einstein's "bold, not to say reckless, hypothesis."

Applying the photon concept to the photoelectric effect, Einstein wrote

$$h\nu = E_0 + K_{\text{max}} \tag{47-11}$$

where $h\nu$ is the energy of the photon. Equation 47-11 says that a photon carries an energy $h\nu$ into the surface. Part of this energy (E_0) is used in causing the electron to pass through the metal surface. The excess energy ($h\nu - E_0$) is given to the electron in the form of kinetic energy; if the electron does not lose energy by internal collisions as it escapes from the metal, it will exhibit it all as kinetic energy after it emerges. Thus K_{max} represents the *maximum* kinetic energy that the photoelectron can have outside the surface; in nearly all cases it will have less energy than this because of internal losses.

Consider how Einstein's photon hypothesis meets the three objections raised against the wave-theory interpretation of the photoelectric effect. As for objection 1 (the lack of dependence of K_{max} on the intensity of illumination), there is complete agreement of the photon theory with experiment. Doubling the light intensity merely doubles the number of photons and thus doubles the photoelectric current; it does not change the energy ($= h\nu$) of the individual photons or the nature of the individual photoelectric processes described by Eq. 47-11.

Objection 2 (the existence of a cutoff frequency) follows from Eq. 47-11. If K_{max} equals zero, we have

$$h\nu_0 = E_0,$$

which asserts that the photon has just enough energy to eject the photoelectrons and none extra to appear as kinetic energy. This quantity E_0 is called the work function of the substance. If ν is reduced below ν_0, the individual photons, no matter how many of them there are (that is, no matter how intense the illumination), will not have enough energy to eject photoelectrons.

Objection 3 (the absence of a time lag) follows from the photon theory because the required energy is supplied in a concentrated bundle. It is *not* spread uniformly over a large area, as in the wave theory.

Although the photon hypothesis certainly fits the facts of photoelectricity, it seems to be in direct conflict with the wave theory of light which, as we have seen in earlier chapters, has been verified in many experiments. Our modern view of the nature of light is that *it has a dual character, behaving like a wave under some circumstances and like a particle, or photon, under others.* We discuss the wave-particle duality at length in Chapter 48. Meanwhile, let us continue our studies of the firm experimental foundation on which the photon concept rests.

Let us rewrite Einstein's photoelectric equation (Eq. 47–11) by substituting eV_0 for $K_{\max}$ (see Eq. 47–9). This yields, after rearrangement,

$$V_0 = \frac{h}{e}\nu - \frac{E_0}{e}. \tag{47–12}$$

Thus Einstein's theory predicts a linear relationship between V_0 and ν, in complete agreement with experiment; see Fig. 47–8. The slope of the experimental curve in this figure should be h/e, or

$$\frac{h}{e} = \frac{ab}{bc} = \frac{2.20\text{ volt} - 0.65\text{ volt}}{(10 \times 10^{14} - 6 \times 10^{14})\text{ cycles/sec}} = 3.9 \times 10^{-15}\text{ volt-sec.}$$

We can find h by multiplying this ratio by the electron charge e,

$$h = (3.9 \times 10^{-15}\text{ volt-sec})(1.6 \times 10^{-19}\text{ coul}) = 6.2 \times 10^{-34}\text{ joule-sec.}$$

From a more careful analysis of this and other data, including data taken with lithium surfaces, Millikan found the value $h = 6.57 \times 10^{-34}$ joule-sec, with an accuracy of about 0.5%. This agreement with the value of h derived from Planck's radiation formula is a striking confirmation of Einstein's photon concept.

▶ **Example 3.** Deduce the work function for sodium from Fig. 47–8.

The intersection of the straight line in Fig. 47–8 with the horizontal axis is the cutoff frequency ν_0. Substituting these values yields

$$\begin{aligned} E_0 = h\nu_0 &= (6.63 \times 10^{-34}\text{ joule-sec})(4.39 \times 10^{14}\text{ cycles/sec}) \\ &= 2.92 \times 10^{-19}\text{ joule} = 1.82\text{ ev.} \end{aligned}$$

◀

47–6 The Compton Effect

Compelling confirmation of the concept of the photon as a concentrated bundle of energy was provided in 1923 by A. H. Compton (1892–1962) who earned a Nobel prize for this work in 1927.* Compton allowed a beam of X-rays of sharply defined wavelength λ to fall on a graphite block, as in Fig. 47–9, and he measured, for various angles of scattering, the intensity of the scattered X-rays as a function of their wavelength. Figure 47–10 shows his experimental results. We see that although the incident beam consists essentially of a single wavelength λ the scattered X-rays have intensity peaks at *two* wavelengths; one of them is the same as the incident wavelength, the other, λ′, being larger by an amount Δλ. This so-called *Compton shift* Δλ varies with the angle at which the scattered X-rays are observed.

The presence of a scattered wave of wavelength λ′ cannot be understood if the incident X-rays are regarded as an electromagnetic wave like that of Fig. 39–11. On this picture the incident wave of frequency ν causes electrons in the scattering block to oscillate at that same frequency. These oscillating electrons, like charges surging back and forth in a small radio transmitting

---

* For an historical account of Compton's researches read "The Scattering of X Rays as Particles," A. H. Compton, p. 817, *Am. J. Phys.*, December 1961.

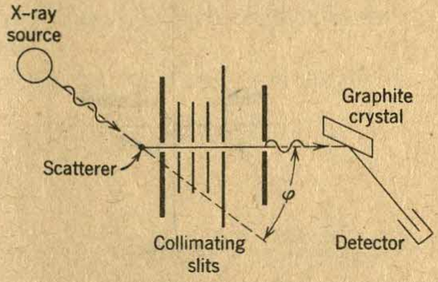


Fig. 47-9 Compton's experimental arrangement. Monochromatic X-rays of wavelength λ fall on a graphite scatterer. The distribution of intensity with wavelength is measured for X-rays scattered at any selected angle φ. The scattered wavelengths are measured by observing Bragg reflections from a crystal; see Eq. 45-8. Their intensities are measured by a detector, such as an ionization chamber.

antenna, radiate electromagnetic waves that again have this same frequency ν. Thus, on the wave picture the scattered wave should have the same frequency ν and the same wavelength λ as the incident wave.

Compton * was able to explain his experimental results by postulating that the incoming X-ray beam was not a wave but an assembly of photons of energy E ($= h\nu$) and that these photons experienced billiard-ball-like collisions with the free electrons in the scattering block. The "recoil" photons emerging from the block constitute, on this view, the scattered radiation. Since the incident photon transfers some of its energy to the electron with

---

* P. W. Debye simultaneously and independently offered the same interpretation.

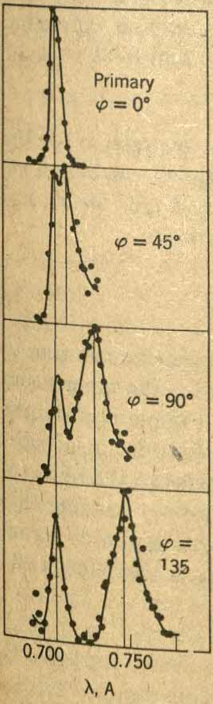


Fig. 47-10 Compton's experimental results. The solid vertical line on the left corresponds to the wavelength λ, that on the right to λ'. Results are shown for four different angles of scattering φ. Note that the Compton shift $\Delta\lambda$ for $\varphi = 90°$ is $h/m_0c = 0.242$ A.

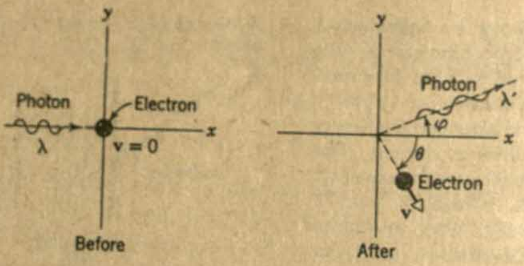


Fig. 47–11 A photon of wavelength λ is incident on an electron at rest. On collision, the photon is scattered at an angle φ with increased wavelength λ′, while the electron moves off with speed v in direction θ.

which it collides, the scattered photon must have a lower energy E'; it must therefore have a lower frequency $\nu'(= E'/h)$, which implies a larger wavelength $\lambda'(= c/\nu')$. This point of view accounts, at least qualitatively, for the wavelength shift $\Delta\lambda$. Notice how different this particle model of X-ray scattering is from that based on the wave picture. Now let us analyze a single photon-electron collision quantitatively.

Figure 47–11 shows a collision between a photon and an electron, the electron assumed to be initially at rest and essentially free, that is, not bound to the atoms of the scatterer. Let us apply the law of conservation of energy to this collision. Since the recoil electrons may have a speed v that is comparable with that of light, we must use the relativistic expression for the kinetic energy of the electron. From Eqs. 47–10 and 8–21 we may write

$$h\nu = h\nu' + (m - m_0)c^2,$$

in which the second term on the right is the relativistic expression for the kinetic energy of the recoiling electron, m being the relativistic mass and m_0 the rest mass of that particle. Substituting c/λ for ν (and c/λ' for ν') and using Eq. 8–20 to eliminate the relativistic mass m leads us to

$$\frac{hc}{\lambda} = \frac{hc}{\lambda'} + m_0c^2\left(\frac{1}{\sqrt{1 - (v/c)^2}} - 1\right). \tag{47–13}$$

Now let us apply the (vector) law of conservation of linear momentum to the collision of Fig. 47–11. We first need an expression for the momentum of a photon. In Section 40–2 we saw that if an object completely absorbs an energy U from a parallel light beam that falls on it the light beam, according to the wave theory of light, will simultaneously transfer to the object a linear momentum given by U/c. On the photon picture we imagine this momentum to be carried along by the individual photons, each photon transporting linear momentum in amount $p = h\nu/c$, where $h\nu$ is the photon energy. Thus, if we substitute λ for c/ν, we can write

$$p = \frac{E}{c} = \frac{h\nu}{c} = \frac{h}{\lambda}. \tag{47–14}$$

This conclusion, that the momentum of a photon is given by h/λ, may also be deduced from the theory of relativity.

For the electron, the relativistic expression for the linear momentum is given by Eq. 9–13, or

$$\mathbf{p}_e = \frac{m_0\mathbf{v}}{\sqrt{1-(v/c)^2}}.$$

We can then write for the conservation of the x component of linear momentum

$$\frac{h}{\lambda} = \frac{h}{\lambda'}\cos\varphi + \frac{m_0 v}{\sqrt{1-(v/c)^2}}\cos\theta \qquad (47\text{–}15)$$

and for the y component

$$0 = \frac{h}{\lambda'}\sin\varphi - \frac{m_0 v}{\sqrt{1-(v/c)^2}}\sin\theta. \qquad (47\text{–}16)$$

Our immediate aim is to find $\Delta\lambda$ ($= \lambda' - \lambda$), the wavelength shift of the scattered photons, so that we may compare it with the experimental results of Fig. 47–10. Compton's experiment did not involve observations of the recoil electron in the scattering block. Of the five collision variables (λ, λ', v, φ, and θ) that appear in the three equations (47–13, 47–15, and 47–16) we may eliminate two. We chose to eliminate v and θ, which deal only with the electron, thereby reducing the three equations to a single relation among the variables.

Carrying out the necessary algebraic steps (see Problem 15) leads to this simple result:

$$\Delta\lambda\ (= \lambda' - \lambda) = \frac{h}{m_0 c}(1 - \cos\varphi). \qquad (47\text{–}17)$$

Thus the Compton shift $\Delta\lambda$ depends only on the scattering angle φ and *not* on the initial wavelength λ. Equation 47–17 predicts within experimental error the experimentally observed Compton shifts of Fig. 47–10. Note from the equation that $\Delta\lambda$ varies from zero (for $\varphi = 0$, corresponding to a "grazing" collision in Fig. 47–11, the incident photon being scarely deflected) to $2h/m_0c$ (for $\varphi = 180°$, corresponding to a "head-on" collision, the incident photon being reversed in direction).

It remains to explain the presence of the peak in Fig. 47–10 for which the wavelength does *not* change on scattering. This peak can be understood as resulting from a collision between a photon and electrons bound in an ionic core in the scattering block. During photon collisions the bound electrons behave like the free electrons that we considered in Fig. 47–11, with the exception that their effective mass is much greater. This is because the ionic core as a whole recoils during the collision. The effective mass M for a carbon scatterer is approximately the mass of a carbon nucleus. Since this nucleus contains 6 protons and 6 neutrons, we have approximately that $M = 12 \times 1840m_0 = 22{,}000m_0$. If we replace m_0 by M in Eq. 47–17, we see that the Compton shift for collisions with tightly bound electrons is immeasurably small.

As in the cavity radiation problem (see Eq. 47–7) and the photoelectric effect (see Eq. 47–11), Planck's constant h is centrally involved in the Compton effect. The quantity h is the central constant of quantum physics. In a universe in which $h = 0$ there would be no quantum physics and classical physics would be valid in the subatomic domain. In particular, as Eq. 47–17 shows, there would be no Compton effect (that is, $\Delta\lambda = 0$) in such a universe.

▶ **Example 4.** X-rays with $\lambda = 1.00$ A are scattered from a carbon block. The scattered radiation is viewed at 90° to the incident beam. (*a*) What is the Compton shift $\Delta\lambda$? (*b*) What kinetic energy is imparted to the recoiling electron?

(*a*) Putting $\varphi = 90°$ in Eq. 47–17, we have, for the Compton shift,

$$\begin{aligned}\Delta\lambda &= \frac{h}{m_0c}(1 - \cos\varphi)\\ &= \frac{6.63 \times 10^{-34}\text{ joule-sec}}{(9.11 \times 10^{-31}\text{ kg})(3.00 \times 10^{8}\text{ meters/sec})}(1 - \cos 90°)\\ &= 2.43 \times 10^{-12}\text{ meter} = 0.0243\text{ A}.\end{aligned}$$

(*b*) If we put K for the kinetic energy of the electron, we can write Eq. 47–13 as

$$\frac{hc}{\lambda} = \frac{hc}{\lambda'} + K.$$

Since $\lambda' = \lambda + \Delta\lambda$, we obtain

$$\frac{hc}{\lambda} = \frac{hc}{\lambda + \Delta\lambda} + K,$$

which reduces to

$$\begin{aligned}K &= \frac{hc\,\Delta\lambda}{\lambda(\lambda + \Delta\lambda)}\\ &= \frac{(6.63 \times 10^{-34}\text{ joule-sec})(3.00 \times 10^{8}\text{ meters/sec})(2.43 \times 10^{-12}\text{ meter})}{(1.00 \times 10^{-10}\text{ meter})(1.00 + 0.024) \times 10^{-10}\text{ meter}}\\ &= 4.73 \times 10^{-17}\text{ joule} = 295\text{ ev}.\end{aligned}$$

The student may show that the initial photon energy E in this case ($= h\nu = hc/\lambda$) is 12,400 ev so that the photon lost about 2.3% of its energy in this collision. A photon whose energy was ten times as large ($= 124{,}000$ ev) can be shown to lose 23% of its energy in a similar collision. This follows from the fact that $\Delta\lambda$ does not depend on the initial wavelength. Hence more energetic X-rays, which have smaller wavelengths, will experience a larger *per cent* increase in wavelength and thus a larger per cent loss in energy. ◀

47–7 Line Spectra

We have seen how Planck successfully explained the nature of the radiation from heated solid objects of which the cavity radiator formed the prototype. Such radiations form *continuous spectra* and are contrasted with *line spectra* such as that of Fig. 47–12, which shows the radiation emitted from iron ions

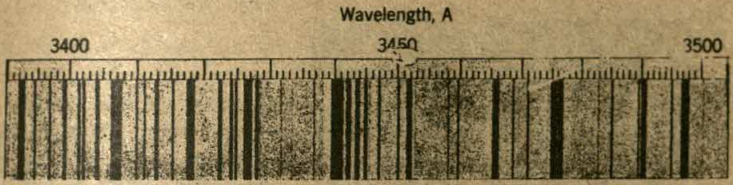


Fig. 47-12 A small portion of the spectrum of iron, in the region 3400 to 3500 A.

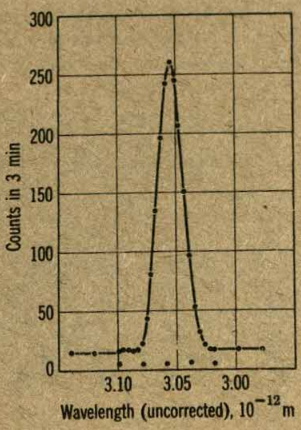
and atoms in an electric arc struck between iron electrodes. We shall see that Planck's quantization ideas, suitably extended, lead to an understanding of line spectra also. The prototype for the study of line spectra is that of atomic hydrogen; being the simplest atom it has the simplest spectrum.

Line spectra are common in all parts of the electromagnetic spectrum. Figure 47-13 shows a spectrum of the γ-rays ($\lambda \cong 10^{-12}$ meter) emitted from a particular radioactive nucleus, an isotope of mercury. Figure 47-14 shows a spectrum of X-rays ($\lambda \cong 10^{-10}$ meter) emitted from a molybdenum target when struck by a 35-kev electron beam. The sharp emission lines are superimposed on a continuous background.

Figure 47-15 shows a spectrum associated with the molecule HCl. It occurs in the infrared, with $\lambda \cong 10^{-6}$ meter. This is an *absorption* spectrum rather than an emission spectrum, as in Fig. 47-12. Experiment shows that isolated atoms and molecules absorb radiation, as well as emit it, at discrete wavelengths.

Figure 47-16 shows a portion of the absorption spectrum of ammonia (NH_3) in the microwave region ($\lambda \cong 10^{-2}$ meter). Finally, Fig. 47-17 shows how radiation in the radio-frequency region ($\lambda \cong 43$ meters) is absorbed by hydrogen molecules placed in a magnetic field.

Fig. 47-13 A wavelength plot for a gamma ray emitted by the nucleus Hg^{198}. (From data by DuMond and co-workers.)



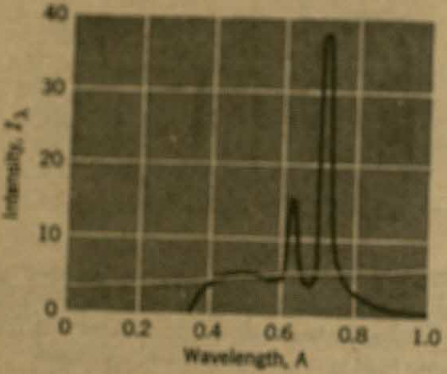


Fig. 47–14 X-rays from a molybdenum target struck by 35-kev electrons. Note the two sharp lines rising above a broad continuous base. The wavelength of the most intense line is 7.1×10^{-11} meter or 0.71 Å. (From data by Ulrey.)

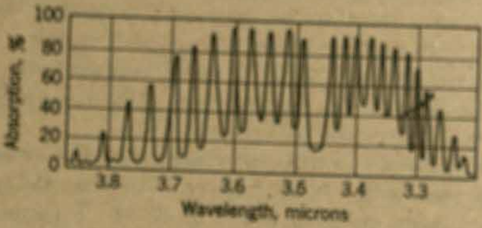
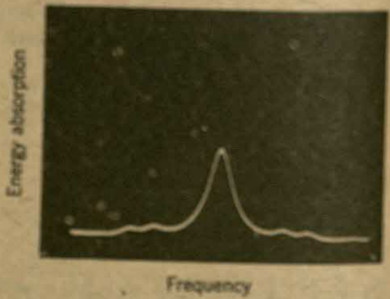


Fig. 47–15 An absorption spectrum of the HCl molecule near $\lambda = 3.5 \times 10^{-6}$ meter = 3.5 μ. (From data by E. S. Imes.)

Fig. 47–16 An oscilloscope trace showing one strong line and four weak lines in the absorption spectrum of ammonia at microwave frequencies.



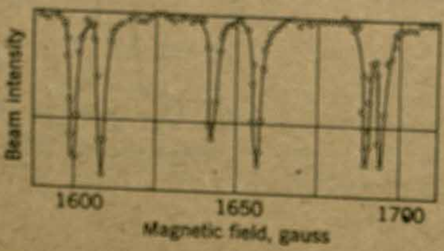


Fig. 47–17 A portion of the absorption spectrum of the protons in molecular hydrogen at $\lambda \cong 43$ meters. In this technique the frequency is left fixed and the sample is placed in a magnetic field, which is varied to scan the spectrum. (From data by Kellogg, Rabi, and Zacharias.)

47-8 The Hydrogen Atom

Figure 47-18 shows the spectrum of hydrogen. A theoretical interpretation of the hydrogen spectrum, based on quantum ideas, was first given by Niels Bohr in 1913.

We might wish to associate the frequency of an emitted spectrum line with the frequency of an electron revolving in an orbit inside the atom. Classical electromagnetism predicts that charges will radiate energy when they are accelerated. In this way electromagnetic waves are emitted from a radio transmitting antenna in which electrons are caused to surge back and forth. This radiation represents a loss of energy for the moving electrons which, in a radio antenna, is compensated for by supplying energy from an oscillator. In an isolated atom, however, no energy is supplied from external sources. We would expect the frequency of the electron and thus that of the emitted radiation to change continuously as the energy drains away. This prediction of classical theory cannot be reconciled with the existence of sharp spectrum lines. Thus classical physics cannot explain the hydrogen, or any other, spectrum.

Bohr circumvented this difficulty by assuming that, like Planck's oscillators, the hydrogen atom exists in certain *stationary states* in which it does not radiate. Radiation occurs *only* when the atom makes a transition from one state, with energy E_k, to a state with lower energy E_j. In equation form

$$h\nu = E_k - E_j, \tag{47-18}$$

where $h\nu$ is the quantum of energy carried away by the photon that is emitted from the atom during the transition.

To learn the allowed frequencies predicted by Eq. 47-18, it is necessary to know the energies of the various stationary states in which a hydrogen atom can exist. This calculation was first carried out by Bohr on the basis of a specific model for the hydrogen atom put forward by him. Bohr's model was highly successful for hydrogen and had a tremendous influence on the

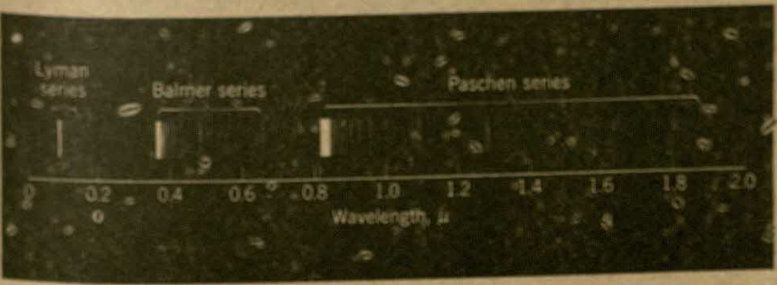


Fig. 47-18 The spectrum of hydrogen. It consists of a number of series of lines, three of which are shown. Within each series the spectrum lines follow a regular pattern, approaching a so-called *series limit* at the short-wave end of the series.

further development of the subject; it is now regarded as an important preliminary stage in the development of a more complete theory of quantum physics.

Let us assume that the electron in the hydrogen atom moves in a circular orbit of radius r centered on its nucleus. We assume that the nucleus, which is a single proton, is so massive that the center of mass of the system is essentially at the position of the proton. Let us calculate the energy E of such an atom.

Writing Newton's second law for the motion of the electron, we have (using Coulomb's law)

$$F = ma,$$

or

$$\frac{e^2}{4\pi\epsilon_0 r^2} = m\frac{v^2}{r}.$$

This allows us to calculate the kinetic energy of the electron, which is

$$K = \tfrac{1}{2}mv^2 = \frac{e^2}{8\pi\epsilon_0 r}. \tag{47-19}$$

The potential energy U of the proton-electron system is given by

$$U = V(-e) = -\frac{e^2}{4\pi\epsilon_0 r}, \tag{47-20}$$

where V $(= e/4\pi\epsilon_0 r)$ is the potential of the proton at the radius of the electron.

The total energy E of the system is

$$E = K + U = -\frac{e^2}{8\pi\epsilon_0 r}. \tag{47-21}$$

Since the orbit radius can apparently take on any value, so can the energy E. The problem of quantizing E reduces to that of quantizing r.

Every property of the orbit is fixed if the radius is given. Equations 47–19, 47–20, and 47–21 show this specifically for the energies K, U, and E. From Eq. 47–19 we can show that the *linear speed* v for the electron is also given in terms of r by

$$v = \sqrt{\frac{e^2}{4\pi\epsilon_0 mr}}. \tag{47-22}$$

The *rotational frequency* ν_0 follows at once from

$$\nu_0 = \frac{v}{2\pi r} = \sqrt{\frac{e^2}{16\pi^3\epsilon_0 mr^3}}. \tag{47-23}$$

The *linear momentum* p follows from Eq. 47–22:

$$p = mv = \sqrt{\frac{me^2}{4\pi\epsilon_0 r}}. \tag{47–24}$$

The *angular momentum* L is given by

$$L = pr = \sqrt{\frac{me^2 r}{4\pi\epsilon_0}}. \tag{47–25}$$

Thus if r is known, the orbit parameters K, U, E, v, ν_o, p, and L are also known. If any one of these quantities is quantized, all of them must be.

At this stage Bohr had no rules to guide him and so made (after some indirect reasoning which we do not reproduce) a bold hypothesis, namely, that the necessary quantization of the orbit parameters shows up most simply when applied to the angular momentum and that, specifically, L can take on only values given by

$$L = n\frac{h}{2\pi} \qquad n = 1, 2, 3, \ldots. \tag{47–26}$$

Planck's constant appears again in a fundamental way; the integer n is a *quantum number*.

Combining Eqs. 47–25 and 47–26 leads to

$$r = n^2 \frac{h^2\epsilon_0}{\pi m e^2} \qquad n = 1, 2, 3, \ldots, \tag{47–27}$$

which tells how r is quantized. Substituting Eq. 47–27 into Eq. 47–21 produces

$$E = -\frac{me^4}{8\epsilon_0{}^2 h^2}\frac{1}{n^2} \qquad n = 1, 2, 3, \ldots, \tag{47–28}$$

which gives directly the energy values of the allowed stationary states.

Figure 47–19 shows the energies of the stationary states and their associated quantum numbers. Equation 47–27 shows that the orbit radius increases as n^2. The upper level in Fig. 47–19, marked $n = \infty$, corresponds to a state in which the electron is completely removed from the atom (that is, $E = 0$ and $r = \infty$). Figure 47–19 also shows some of the quantum jumps that take place between the different stationary states.

Combining Eqs. 47–18 and 47–28 allows us to write a completely theoretical formula for the frequencies of the lines in the hydrogen spectrum. It is

$$\nu = \frac{me^4}{8\epsilon_0{}^2 h^3}\left(\frac{1}{j^2} - \frac{1}{k^2}\right) \tag{47–29}$$

in which j and k are integers describing, respectively, the lower and the upper stationary states. The corresponding wavelengths can easily be

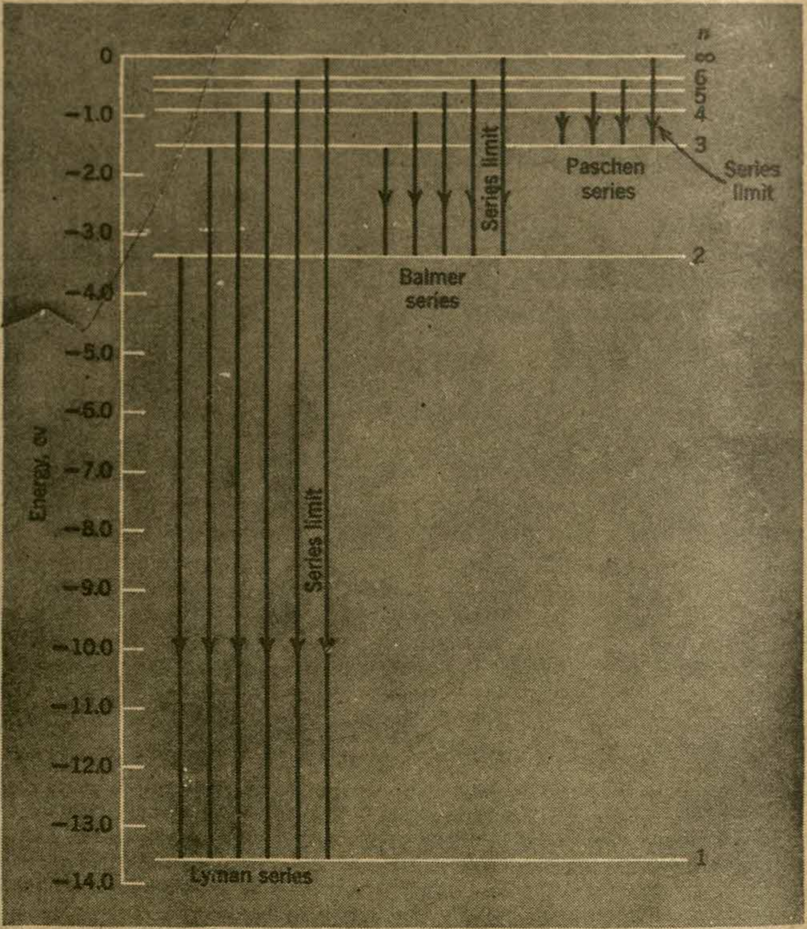


Fig. 47–19 An energy level diagram for hydrogen showing the quantum number n for each level and some of the transitions that appear in the spectrum. An infinite number of levels is crowded in between the levels marked $n = 6$ and $n = \infty$. Compare this figure carefully with Fig. 47–18.

found from $\lambda = c/\nu$. Table 47-1 shows some wavelengths so calculated; it should be compared carefully with Figs. 47-18 and 47-19.

Table 47-1

THE HYDROGEN SPECTRUM

(Some selected lines)

Name of series	Quantum Number		Wavelength, A
	j (lower state)	k (upper state)	
Lyman	1	2	1216
	1	3	1026
	1	4	970
	1	5	949
	1	6	940
	1	∞	912
Balmer	2	3	6563
	2	4	4861
	2	5	4341
	2	6	4102
	2	7	3970
	2	∞	3650
Paschen	3	4	18751
	3	5	12818
	3	6	10938
	3	7	10050
	3	8	9546
	3	∞	8220

▶ **Example 5.** Calculate the binding energy of the hydrogen atom (the energy binding the electron to the nucleus) from Eq. 47-28.

The binding energy is numerically equal to the energy of the lowest state in Fig. 47-19. The largest negative value of E in Eq. 47-28 is found for $n = 1$. This yields

$$E = -\frac{me^4}{8\epsilon_0{}^2h^2}$$

$$= -\frac{(9.11 \times 10^{-31}\text{ kg})(1.60 \times 10^{-19}\text{ coul})^4}{(8)(8.85 \times 10^{-12}\text{ coul}^2/\text{nt-m}^2)^2(6.63 \times 10^{-34}\text{ joule-sec})^2}$$

$$= -2.17 \times 10^{-18}\text{ joule} = -13.6\text{ ev},$$

in which agrees with the experimentally observed binding energy for hydrogen. ◀

47-9 The Correspondence Principle

Although all theories in physics have limitations, they usually do not break down abruptly but in a continuous way, yielding results that agree less and less well with experiment. Thus the predictions of Newtonian mechanics become less and less accurate as the speed is made to approach that of light. A similar relationship must exist between quantum physics and classical physics; it remains to find the circumstances under which the latter theory is revealed as a special case of the former.

The radius of the lowest energy state in hydrogen (the so-called *ground state*) is found by putting $n = 1$ in Eq. 47–27; it turns out to be 5.3×10^{-11} meter. If $n = 10{,}000$, however, the radius is $(10{,}000)^2$ times as large or 5.3 mm. This "atom" is so large that we suspect that its behavior should be accurately described by classical physics. Let us test this by computing the frequency of the emitted light on the basis of both classical and quantum assumptions. These calculations should differ at small quantum numbers but should agree at large quantum numbers. The fact that *quantum physics reduces to classical physics at large quantum numbers* is called the *correspondence principle*. This principle, credited to Niels Bohr, was very useful during the years in which quantum physics was being developed. Bohr, in fact, based his theory of the hydrogen atom on correspondence principle arguments.

Classically, the frequency of the light emitted from an atom is equal to ν_0, its frequency of revolution * in its orbit. This can be expressed in terms of a quantum number n by combining Eqs. 47–23 and 47–27 to obtain

$$\nu_0 = \frac{me^4}{8\epsilon_0{}^2h^3}\frac{2}{n^3}. \tag{47–30}$$

Quantum physics predicts that the frequency ν of the emitted light is given by Eq. 47–29. Considering a transition between an orbit with quantum number $k = n$ and one with $j = n - 1$ leads to

$$\begin{aligned}\nu &= \frac{me^4}{8\epsilon_0{}^2h^3}\left[\frac{1}{(n-1)^2} - \frac{1}{n^2}\right]\\ &= \frac{me^4}{8\epsilon_0{}^2h^3}\left[\frac{2n-1}{(n-1)^2n^2}\right].\end{aligned} \tag{47–31}$$

As $n \to \infty$ the expression in the square brackets above approaches $2/n^3$ so that $\nu \to \nu_0$ as $n \to \infty$. Table 47–2 illustrates this example of the correspondence principle.

---

* Integral multiples of the frequency also exist but may be ignored without affecting the present argument.

Table 47–2

THE CORRESPONDENCE PRINCIPLE AS APPLIED TO THE HYDROGEN ATOM

Quantum Number, n	Frequency of Revolution in Orbit (Eq. 47–30) cycles/sec	Frequency of Transition to Next Lowest State (Eq. 47–31) cycles/sec	Difference, %
2	8.20×10^{14}	24.6×10^{14}	67
5	5.26×10^{13}	7.38×10^{13}	29
10	6.57×10^{12}	7.72×10^{12}	14
50	5.25×10^{10}	5.42×10^{10}	3
100	6.578×10^{9}	6.677×10^{9}	1.5
1,000	6.5779×10^{6}	6.5878×10^{6}	0.15
10,000	6.5779×10^{3}	6.5789×10^{3}	0.015

QUESTIONS

1. "Pockets" formed by the coals in a coal fire seem brighter than the coals themselves. Is the temperature in such pockets appreciably higher than the surface temperature of an exposed glowing coal?

2. The relation $R = \sigma T^4$ (Eq. 47–2) is exact for true cavities and holds for all temperatures. Why don't we use this relation as the basis of a definition of temperature at, say, 100°C?

3. Do all incandescent solids obey the fourth-power law of temperature, as Eq. 47–3 seems to suggest?

4. A hole in the wall of a cavity radiator is sometimes called a *black body*. Why?

5. It is stated that if we look into a cavity whose walls are maintained at a constant temperature no details of the interior are visible. Does this seem reasonable?

6. How can a *photon* energy be given by $E = h\nu$ (Eq. 47–10) when the very presence of the frequency ν in the formula implies that light is a *wave*?

7. In the photoelectric effect, why does the existence of a cutoff frequency speak in favor of the photon theory and against the wave theory?

8. Why are photoelectric measurements so sensitive to the nature of the photoelectric surface?

9. Does Einstein's theory of photoelectricity, in which light is postulated to be a photon, invalidate Young's interference experiment?

10. List and discuss carefully the assumptions made by Planck in connection with the cavity radiation problem, by Einstein in connection with the photoelectric effect, and by Bohr in connection with the hydrogen atom problem.

11. In Bohr's theory for the hydrogen atom orbits, what is the implication of the fact that the potential energy is negative and is greater in magnitude than the kinetic energy?

12. Can a hydrogen atom absorb a photon whose energy exceeds its binding energy (13.6 ev)?

13. Discuss Example 1 in terms of the correspondence principle.

14. According to classical mechanics, an electron moving in an orbit should be able to do so with any angular momentum whatever. According to Bohr's theory of the hydrogen atom, however, the angular momentum is quantized according to $L = nh/2\pi$. Reconcile these two statements, using the correspondence principle.

PROBLEMS

1. At what wavelength does a cavity radiator at 6000°K radiate most per unit wavelength? Solve either analytically or graphically.

2. Using Fig. 47–1, verify that the emissivity of tungsten at 2000°K is 0.259.

3. A cavity radiator at 6000°K has a hole 0.10 mm in diameter drilled in its wall. At what rate do photons in the range 5500 to 5510 A escape from this hole?

4. A cavity whose walls are held at 4000°K has a circular aperture 5.0 mm in diameter. (*a*) At what rate does energy in the visible range (defined to extend from 0.40 to 0.70 μ) escape from this hole? (*b*) What fraction of the total radiation escaping from the cavity does this represent? Solve either analytically or graphically.

5. Show that Wien's law (p. 1085) is a special case of Planck's law (Eq. 47–6) for short wavelengths or low temperatures.

6. Solar radiation falls on the earth at a rate of 2.0 cal/cm^2-min. How many photons/cm^2-min is this, assuming an average wavelength of 5500 A?

7. A spectral emission line, important in radioastronomy, has a wavelength of 21 cm. To what photon energy does this correspond?

8. The energy required to remove an electron from sodium is 2.3 ev. Does sodium show a photoelectric effect for orange light, with $\lambda = 6800$ A?

9. Light of a wavelength 2000 A falls on an aluminum surface. In aluminum 4.2 ev are required to remove an electron. What is the kinetic energy of (*a*) the fastest and (*b*) the slowest emitted photoelectrons? (*c*) What is the stopping potential? (*d*) What is the cutoff wavelength for aluminum?

10. The work function for a clean lithium surface is 2.3 ev. Make a rough plot of the stopping potential V_0 versus the frequency of the incident light for such a surface.

11. In Example 2 suppose that the "target" is a single gas atom of 1.0 A radius and that the intensity of the light source is reduced to 10^{-5} watt. If the binding energy of the most loosely bound electron in the atom is 20 ev, what time lag for the photoelectric effect is expected on the basis of the wave theory of light?

12. A 100-watt sodium vapor lamp radiates uniformly in all directions. (*a*) At what distance from the lamp will the average density of photons be 10/cm^3? (*b*) What is the average density of photons 2.0 meters from the lamp? Assume the light to be monochromatic, with $\lambda = 5890$ A.

13. Show, by analyzing a collision between a photon and a free electron (using relativistic mechanics), that it is impossible for a photon to give *all* of its energy to the free electron. In other terms, the photoelectric effect cannot occur for completely free electrons; the electrons must be bound in a solid or in an atom.

14. Calculate the per cent change in photon energy for a Compton collision with φ in Fig. 47–11 equal to 90° for radiation in (*a*) the microwave range, with $\lambda = 3.0$ cm, (*b*) the visible range, with $\lambda = 5000$ A, (*c*) the X-ray range, with $\lambda = 1.00$ A, and (*d*) the gamma ray range, the energy of the gamma-ray photons being 1.0 Mev. What are your conclusions about the importance of the Compton effect in these various regions of the electromagnetic spectrum, judged solely by the criterion of energy loss in a single Compton encounter?

15. Carry out the necessary algebra to eliminate v and θ from Eqs. 47–13, 47–15, and 47–16 to obtain the Compton shift relation (Eq. 47–17).

16. (*a*) Using Bohr's formula, calculate the three longest wavelengths in the Balmer series. (*b*) Between what wavelength limits does the Balmer series lie?

17. In the ground state of the hydrogen atom, according to Bohr's theory, what are (*a*) the quantum number, (*b*) the orbit radius, (*c*) the angular momentum, (*d*) the linear momentum, (*e*) the angular velocity, (*f*) the linear speed, (*g*) the force on the electron, (*h*) the acceleration of the electron, (*i*) the kinetic energy, (*j*) the potential energy, and (*k*) the total energy?

18. How do the quantities (*b*) to (*k*) in Problem 17 vary with the quantum number?

19. How much energy is required to remove an electron from a hydrogen atom in a state with $n = 8$?

20. A hydrogen atom is excited from a state with $n = 1$ to one with $n = 4$. (*a*) Calculate the energy that must be absorbed by the atom. (*b*) Calculate and display on an energy-level diagram the different photon energies that may be emitted if the atom returns to its $n = 1$ state. (*c*) Calculate the recoil speed of the hydrogen atom, assumed initially at rest, if it makes the transition from $n = 4$ to $n = 1$ in a single quantum jump.

21. A hydrogen atom in a state having a *binding energy* (this is the energy required to remove an electron) of 0.85 ev makes a transition to a state with an *excitation energy* (this is the difference in energy between the state and the ground state) of 10.2 ev. (*a*) Find the energy of the emitted photon. (*b*) Show this transition on an energy-level diagram for hydrogen, labeling the appropriate quantum numbers.

22. Show on an energy-level diagram for hydrogen the quantum numbers corresponding to a transition in which the wavelength of the emitted photon is 1216 A.

23. A neutron, with kinetic energy of 6.0 ev, collides with a resting hydrogen atom in its ground state. Apply the laws of conservation of momentum and of energy to this collision and show that it must be elastic (that is, *kinetic* energy must be conserved).

24. Apply Bohr's theory to singly ionized helium, that is, to a helium atom with one electron removed. What relationship exists between this spectrum and the hydrogen spectrum?

25. Using Bohr's theory, calculate the energy required to remove the electron from singly ionized helium.

26. *Positronium*. Apply Bohr's theory to the positronium atom. This consists of a positive and a negative electron revolving around their center of mass, which lies halfway between them. (*a*) What relationship exists between this spectrum and the hydrogen spectrum? (*b*) What is the radius of the ground state orbit? (Hint: It will be necessary to analyze this problem from first principles because this "atom" has no nucleus; both particles revolve about a point halfway between them.)

27. *Muonic atoms*. Apply Bohr's theory to a muonic atom, which consists of a nucleus of charge Ze with a negative muon (an elementary particle with a charge of $-e$ and a mass m that is 207 times as large as the electron mass; see Appendix *F*) circulating about it. Calculate (*a*) the radius of the first Bohr orbit, (*b*) the ionization energy, and (*c*) the wavelength of the most energetic photon that can be emitted. Assume that the muon is circulating about a hydrogen nucleus ($Z = 1$).

28. If an electron is rotating in an orbit at frequency ν_0, classical electromagnetism predicts that it will radiate energy not only at this frequency but also at $2\nu_0$, $3\nu_0$, $4\nu_0$, etc.; see footnote on p. 1104. Show that this is also predicted by Bohr's theory of the hydrogen atom in the limiting case of large quantum numbers.

29. In Table 47–2 show that the quantity in the last column is given by

$$\frac{100(\nu - \nu_0)}{\nu} \cong \frac{150}{n}$$

for large quantum numbers.

30. If the angular momentum of the earth due to its motion around the sun were quantized according to Bohr's relation $L = nh/2\pi$, what would the quantum number be? Could such quantization be detected if it existed?

Waves and Particles

CHAPTER 48

48–1 Matter Waves

In 1924 Louis de Broglie of France reasoned that (*a*) nature is strikingly symmetrical in many ways; (*b*) our observable universe is composed entirely of light and matter; (*c*) if light has a dual, wave-particle nature, perhaps matter has also. Since matter was then regarded as being composed of particles, de Broglie's reasoning suggested that one should search for a wave-like behavior for matter.

De Broglie's suggestion might not have received serious attention had he not predicted what the expected wavelength of the so-called matter waves would be. We recall that about 1680 Huygens put forward a wave theory of light that did not receive general acceptance, in part because Huygens was not able to state what the wavelength of the light was. When Thomas Young rectified this defect in 1800, the wave theory of light started on its way to acceptance.

De Broglie assumed that the wavelength of the predicted matter waves was given by the same relationship that held for light namely, Eq. 47–14, or

$$\lambda = \frac{h}{p}, \tag{48–1}$$

which connects the wavelength of a light wave with the momentum of the associated photons. The dual nature of light shows up strikingly in this equation and also in Eq. 47–10 ($E = h\nu$). Each equation contains within its structure both a wave concept (ν and λ) and a particle concept (E and p) *De Broglie predicted that the wavelength of matter waves would also be given by Eq. 48–1, where p would now be the momentum of the particle of matter.*

▶ **Example 1.** What wavelength is predicted by Eq. 48-1 for a beam of electrons whose kinetic energy is 100 ev?

The velocity of the electrons is found from $K = \frac{1}{2}mv^2$, or

$$v = \sqrt{\frac{2K}{m}} = \sqrt{\frac{(2)(100\text{ ev})(1.6 \times 10^{-19}\text{ joule/ev})}{9.1 \times 10^{-31}\text{ kg}}}$$

$$= 5.9 \times 10^6\text{ meters/sec.}$$

The momentum follows from

$$p = mv = (9.1 \times 10^{-31}\text{ kg})(5.9 \times 10^6\text{ meters/sec}) = 5.4 \times 10^{-24}\text{ kg-m/sec.}$$

The wavelength (called the *de Broglie wavelength*) is found from Eq. 48-1 or

$$\lambda = \frac{h}{p} = \frac{6.6 \times 10^{-34}\text{ joule-sec}}{5.4 \times 10^{-24}\text{ kg-m/sec}} = 1.2\text{ A.}$$

This is the same order of magnitude as the size of an atom or the spacing between adjacent planes of atoms in a solid. ◀

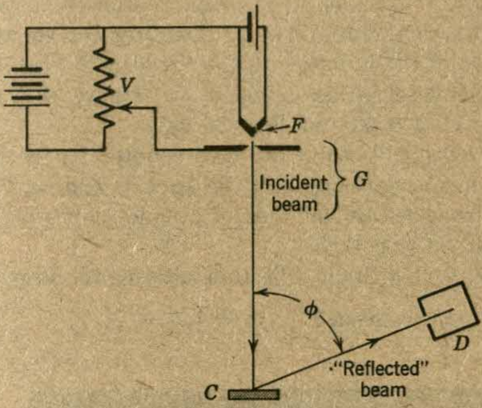


Fig. 48-1 The apparatus of Davisson and Germer. Electrons from filament F are accelerated by a variable potential difference V. After "reflection" from crystal C they are collected by detector D.

In 1926 Elsasser pointed out that the wave nature of matter might be tested in the same way that the wave nature of X-rays was first tested, namely, by allowing a beam of electrons of the appropriate energy to fall on a crystalline solid. The atoms of the crystal serve as a three-dimensional array of diffracting centers for the electron "wave"; we should look for strong diffracted peaks in certain characteristic directions, just as for X-ray diffraction.

This idea was tested by C. J. Davisson and L. H. Germer in this country and by G. P. Thomson in Scotland.* Figure 48-1 shows the apparatus of

---

* For an historical account of Thomson's researches see "Early Work in Electron Diffraction," Sir George Thomson, p. 821, *Am. J Phys.*, December, 1961.

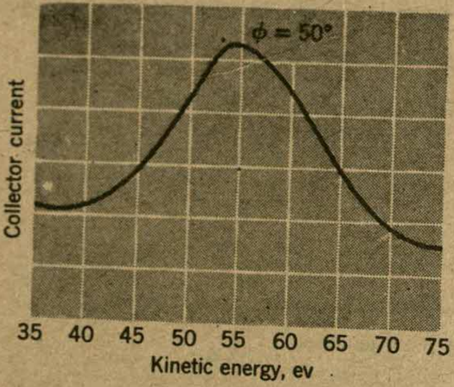


Fig. 48–2 The collector current in detector D in Fig. 48–1 as a function of the kinetic energy of the incident electrons, showing a diffraction maximum. The angle ϕ in Fig. 48–1 is adjusted to 50°. If an appreciably smaller or larger value is used, the diffraction maximum disappears.

Davisson and Germer: Electrons from a heated filament are accelerated by a variable potential difference V and emerge from the "electron gun" G with kinetic energy eV. This electron beam is allowed to fall at normal incidence on a single crystal of nickel at C. Detector D is set at a particular angle ϕ and readings of the intensity of the "reflected" beam are taken at various values of the accelerating potential V. Figure 48–2 shows that a strong beam occurs at $\phi = 50°$ for $V = 54$ volts.

All such strong "reflected" beams can be accounted for by assuming that the electron beam has a wavelength, given by $\lambda = h/p$, and that "Bragg reflections" occur from certain families of atomic planes precisely as described for X-rays in Section 45–5.

Figure 48–3 shows such a Bragg reflection, obeying the Bragg relationship

$$m\lambda = 2d \sin \theta \qquad m = 1, 2, 3, \ldots; \tag{48–2}$$

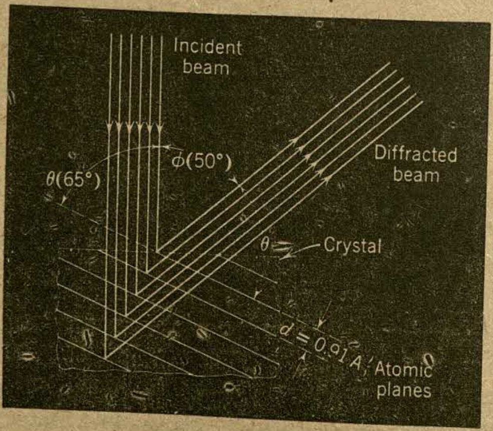


Fig. 48–3 The strong diffracted beam at $\phi = 50°$ and $V = 54$ volts arises from wave-like "reflection" from the family of atomic planes shown, for $d = 0.91$ A. The Bragg angle θ is 65°. For simplicity, refraction of the diffracted wave as it leaves the surface is ignored.

For the conditions of Fig. 48–3 the effective interplanar spacing d can be shown by X-ray analysis to be 0.91 A. Since ϕ equals 50°, it follows that θ equals $90° - \frac{1}{2} \times 50°$ or 65°. The wavelength to be calculated from Eq. 48–2, if we assume $m = 1$, is

$$\lambda = 2d \sin \theta = 2(0.91 \text{ A})(\sin 65°) = 1.65 \text{ A}.$$

The wavelength calculated from the de Broglie relationship $\lambda = h/p$ is, for 54-ev electrons (see Example 1), 1.64 A. This excellent agreement, combined with much similar evidence, is a convincing argument for believing that electrons are wave-like in some circumstances.

Not only electrons but all other particles, charged or uncharged, show wave-like characteristics. Beams of slow neutrons from nuclear reactors are routinely used to investigate the atomic structure of solids. Figure 48–4 shows a "neutron diffraction pattern" for finely powdered lead.

The evidence for the existence of matter waves with wavelengths given by Eq. 48–1 is strong indeed. Nevertheless, the evidence that matter is composed of particles remains equally strong; see Fig. 10–10. Thus, for matter as for light, we must face up to the existence of a dual character; *matter behaves in some circumstances like a particle and in others like a wave.*

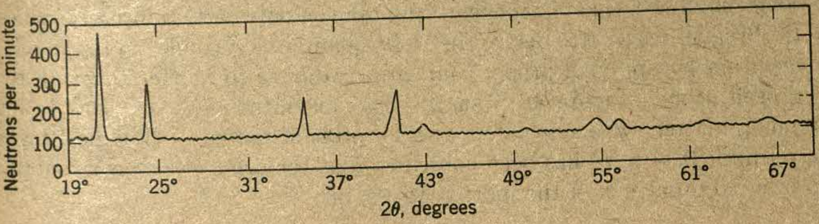


Fig. 48–4 A diffraction pattern for powdered lead, using a monochromatic neutron beam from a nuclear reactor. The peaks represent "Bragg reflections" from the various atomic planes, θ being the corresponding "Bragg angle."

48–2 Atomic Structure and Standing Waves

The motion of electrons in beams is not bounded or limited in the beam direction. We can make an analogy to a sound wave in a long gas-filled tube, a wave traveling down a long string, or an electromagnetic wave in a long waveguide. All four cases can be described by appropriate traveling waves and, significantly, waves of *any* wavelength (within a certain range) can be propagated.

Let these last three waves be *bounded* by imposing physical restrictions. For the sound wave this corresponds to inserting end walls on a section of the long gas-filled pipe, thus forming an acoustic resonant cavity (Section 38–5). For the waves in the string it corresponds to removing a finite section of string and clamping it at each end, as a violin string (Section 20–5).

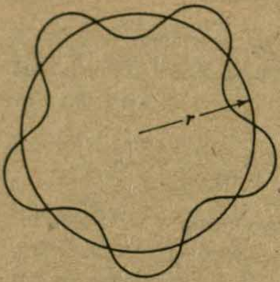


Fig. 48–5 Showing how an electron wave can be adjusted in wavelength to fit an integral number of times around the circumference of a Bohr orbit of radius r. This concept, like the Bohr orbit concept, is now regarded as oversimplified.

For the electromagnetic wave it corresponds to inserting end caps on a finite length of waveguide, thus forming an electromagnetic resonant cavity (Section 38–6).

Two important changes occur: (*a*) the motions are now represented by *standing* rather than traveling waves and (*b*) *only certain wavelengths* (*or frequencies*) *can now exist.* This *quantization* of the wavelength is a direct result of bounding or limiting the wave. We expect that if electrons are limited in their motions by being localized in an atom that (*a*) the electron motion can be represented by a standing matter wave, and (*b*) the electron motion will become quantized, that is, its energy can take on only certain discrete values.

De Broglie was able to derive the Bohr quantization condition for angular momentum by applying proper *boundary conditions* to matter waves in the hydrogen atom. Figure 48–5 suggests an instantaneous "snapshot" of a standing matter wave associated with an orbit of radius r. The de Broglie wavelength ($\lambda = h/p$) has been chosen so that the orbit of radius r contains an integral number n of the matter waves, or

$$\frac{2\pi r}{\lambda} = \frac{2\pi r}{(h/p)} = n \qquad n = 1, 2, 3, \ldots.$$

This leads at once to

$$L = pr = n\frac{h}{2\pi} \qquad n = 1, 2, 3, \ldots,$$

which is the Bohr quantization condition for L.

48–3 Wave Mechanics

The idea that the stationary states in atoms correspond to standing matter waves was taken up by Erwin Schrödinger in 1926 and used by him as the foundation of *wave mechanics*, one of several equivalent formulations of quantum physics.

An important quantity in wave mechanics is the *wave function* Ψ, which measures the "wave disturbance" of matter waves. For waves on strings the "wave disturbance" may be measured by a transverse displacement y; for sound waves it may be measured by a pressure variation p; for electromagnetic waves it may be measured by the electric field vector $\mathbf{E}$.

We make the physical meaning of the wave disturbance Ψ clear in Section 48-4. Meanwhile, let us study the wave function $\Psi(x, t)$ for a simple, one-dimensional problem, that of the possible motions a particle of mass m confined between rigid walls of separation l as in Fig. 48-6*b*. The wave function can be obtained by analogy with a known mechanical problem, that of the natural modes of vibration of a string of length l, clamped at each end as in Fig. 48-6*a*.

In the vibrating string the boundary conditions require that nodes exist at each end. This means that the wavelength λ must be chosen so that

$$l = n\frac{\lambda}{2} \qquad n = 1, 2, 3, \ldots,$$

or that the wavelength λ is "quantized" by the requirement that

$$\lambda = \frac{2l}{n} \qquad n = 1, 2, 3, \ldots. \tag{48-3}$$

The wave disturbance for the string is represented by a standing wave whose equation was shown in Section 19-9 to be

$$y = 2y_m \sin kx \cos \omega t,$$

where $\omega\ (= 2\pi\nu)$ is the angular frequency of the wave and $k\ (= 2\pi/\lambda)$ is the wave number. Since λ is quantized, k must be also, or

$$k = \frac{2\pi}{\lambda} = \frac{n\pi}{l} \qquad n = 1, 2, 3, \ldots.$$

which leads to

$$y = \left[2y_m \sin\frac{n\pi x}{l}\right] \cos \omega t \qquad n = 1, 2, 3, \ldots. \tag{48-4}$$

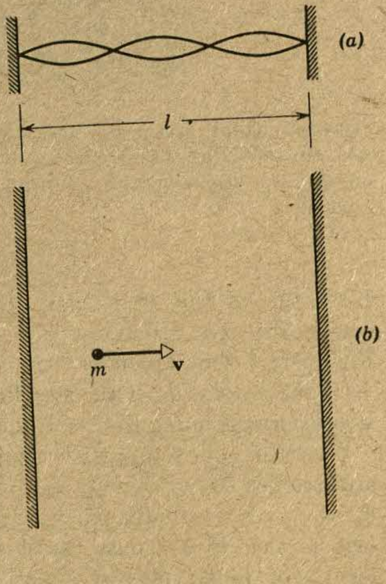


Fig. 48-6 (*a*) A stretched string of length l clamped between rigid supports. (*b*) A particle of mass m and velocity v confined to move between rigid walls a distance l apart.

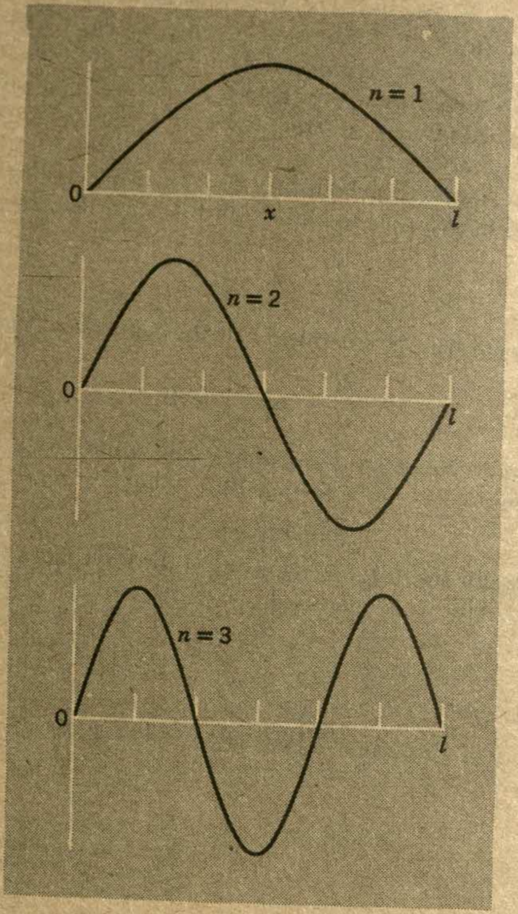


Fig. 48–7 Three quantized modes of vibration for the string of Fig. 48–6*a*. The figure also represents three of the quantized wave functions for the particle of Fig. 48–6*b*. The ordinate is a displacement amplitude in the first case and a wave-function amplitude in the second.

Inspection of Eq. 48–4 shows that no matter what value of n is selected nodes exist at $x = 0$ and at $x = l$, as required by the boundary conditions. Figure 48–7 shows plots of the quantity in the square brackets in this equation (the amplitude of the standing wave) for the modes of vibration of the string corresponding to $n = 1$, 2, and 3.

Consider now the particle confined between rigid walls. Since the walls are assumed to be perfectly rigid, the particle cannot penetrate them so that Ψ, which represents the particle motion in some way not yet clearly specified (see Section 48–4), must vanish at $x = 0$ and $x = l$. The allowed wavelengths of the matter waves must be given by Eq. 48–3, or

$$\lambda = \frac{2l}{n}.$$

Replacing λ by h/p (see Eq. 48–1) leads to

$$p = \frac{nh}{2l}, \tag{48–5}$$

which shows that the *linear momentum* of the particle is quantized. The momentum p ($= mv$) is related to the energy E (which is entirely kinetic and is equal to $\frac{1}{2}mv^2$) by

$$p = \sqrt{2mE}. \tag{48–6}$$

Combining Eqs. 48–5 and 48–6 leads to the quantization condition for E, or

$$E = n^2 \frac{h^2}{8ml^2} \qquad n = 1, 2, 3, \ldots. \tag{48–7}$$

The particle cannot have any energy, as we would expect classically, but only energies given by Eq. 48–7.

The matter wave is described, in strict analogy with Eq. 48–4, by

$$\Psi = \left[\Psi_m \sin \frac{n\pi x}{l}\right] \cos \omega t \qquad n = 1, 2, 3, \ldots. \tag{48–8}$$

Figure 48–7 can serve equally well to show how the amplitude of the standing matter waves for the states of motion corresponding to $n = 1$, 2, and 3 varies throughout the box. We see clearly in this problem how the act of localizing or bounding a particle leads to energy quantization.

▶ **Example 2.** Consider an electron ($m = 9.1 \times 10^{-31}$ kg) confined by electrical forces to move between two rigid "walls" separated by 1.0×10^{-9} meter, which is about five atomic diameters. Find the quantized energy values for the three lowest stationary states.

From Eq. 48–7, for $n = 1$, we have

$$E = n^2 \frac{h^2}{8ml^2} = (1)^2 \frac{(6.6 \times 10^{-34}\ \text{joule-sec})^2}{(8)(9.1 \times 10^{-31}\ \text{kg})(1.0 \times 10^{-9}\ \text{meter})^2}$$

$$= 6.0 \times 10^{-20}\ \text{joule} = 0.38\ \text{ev}.$$

The energies for the next two states ($n = 2$ and $n = 3$) are $2^2 \times 0.38$ ev $= 1.5$ ev and $3^2 \times 0.38$ ev $= 3.4$ ev.

Example 3. Consider a grain of dust ($m = 1.0\ \mu\text{gm} = 1.0 \times 10^{-9}$ kg) confined to move between two rigid walls separated by 0.1 mm ($= 10^{-4}$ meter). Its speed is only 10^{-6} meter/sec, so that it requires 100 sec to cross the gap. What quantum number describes this motion?

The energy is

$$E\ (= K) = \tfrac{1}{2}mv^2 = \tfrac{1}{2}(10^{-9}\ \text{kg})(10^{-6}\ \text{meter/sec})^2$$

$$= 5 \times 10^{-22}\ \text{joule}.$$

Solving Eq. 48-7 for n yields

$$n = \sqrt{8mE}\,\frac{l}{h} = \sqrt{(8)(10^{-9}\text{ kg})(5 \times 10^{-22}\text{ joule})}\left(\frac{10^{-4}\text{ meter}}{6.6 \times 10^{-34}\text{ joule-sec}}\right)$$
$$= 3 \times 10^{14}.$$

Even in these extreme conditions the quantized nature of the motion would never be apparent; we cannot distinguish experimentally between $n = 3 \times 10^{14}$ and $n = 3 \times 10^{14} + 1$. Classical physics, which fails completely for the problem of Example 2, works extremely well for this problem. ◀

48-4 Meaning of Ψ

Max Born first suggested that the quantity Ψ^2 at any particular point is a measure of *the probability that the particle will be near that point.* More exactly, if a volume element dV is constructed at that point, the probability that the particle will be found in the volume element at a given instant is $\Psi^2\,dV$. This interpretation of Ψ provides a statistical connection between the wave and the associated particle; it tells us where the particle is *likely to be*, not where it *is*.

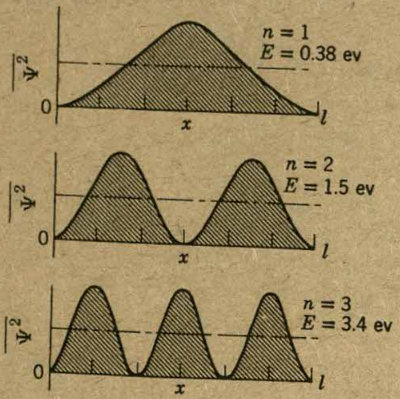


Fig. 48-8 The "probability functions" for three states of motion of the particle of Fig. 48-6*b*, along with the corresponding quantized energies for the conditions of Example 2. The horizontal lines show the predictions of classical mechanics, in which the probability function is constant for all positions of the particle.

For the particle confined between rigid walls the probability that the particle will lie between two planes that are distance x and $x + dx$ from one wall (see Fig. 48-8) is given by

$$\Psi^2\,dx = \Psi_m{}^2 \sin^2 \frac{n\pi}{l} x \cos^2 \omega t\,dx.$$

Since we are more interested in the particle's spatial location than in its time behavior, we average Ψ^2 over one cycle of the motion. This is equivalent to replacing $\cos^2 \omega t$ by its average value for one cycle, namely one-half, or

$$\overline{\Psi^2} = \tfrac{1}{2}\Psi_m{}^2 \sin^2 \frac{n\pi x}{l} \qquad n = 1, 2, 3, \ldots. \tag{48-9}$$

Figure 48-8 shows $\overline{\Psi^2}$ for the three stationary states corresponding to $n = 1$, 2, and 3. Note that for $n = 1$ the particle is more likely to be near the center than the ends. This is in sharp contradiction to the results of classical physics, according to which the particle has the same probability of being located anywhere between the walls, as shown by the horizontal line in Fig. 48–8.

The problem of a particle confined between rigid walls has little real application in physics. We would prefer to illustrate the wave mechanics of Schrödinger by applying it to a more experimentally realizable situation, such as the hydrogen atom. Only mathematical complexity prevents us from doing this. We state without proof that when this problem is solved by wave mechanics the motion of the electron in the ground state of the atom, defined by putting $n = 1$ in Eq. 47–28, is described by the following wave function,

$$\Psi = \frac{1}{\sqrt{\pi a^3}} e^{-r/a} \cos \omega t, \tag{48–10}$$

where

$$a = \frac{h^2 \epsilon_0}{\pi m e^2}.$$

Putting $n = 1$ in Eq. 47–27 shows that a is the radius of the ground-state orbit in Bohr's theory. This special interpretation has little meaning in wave mechanics; a is taken here merely as a convenient unit of length when dealing with atomic problems, having the value 0.529 A.

▶ **Example 4.** Consider two hypothetical spherical shells centered on the nucleus of a hydrogen atom with radii r and $r + dr$. What is the probability $P(r)$ that the electron will lie between these shells, as a function of r?

This probability is $\Psi^2\, dV$, where dV is the volume between the shells, or $4\pi r^2\, dr$. Thus

$$\Psi^2\, dV = \left(\frac{1}{\sqrt{\pi a^3}} e^{-r/a} \cos \omega t\right)^2 (4\pi r^2\, dr) = P(r)\, dr.$$

Averaging over the time (that is, replacing $\cos^2 \omega t$ by $\frac{1}{2}$) yields for the probability $\overline{P(r)}$

$$\overline{P(r)} = \frac{2r^2}{a^3} e^{-2r/a}.$$

Figure 48–9 shows a plot of this function. Note that the most probable location for the electron corresponds to the first Bohr radius. Thus in wave mechanics we do not say that the electron in the $n = 1$ state in hydrogen goes around the nucleus in a circular orbit of 0.529 A radius but only that the electron is more likely to be found at this distance from the nucleus than at any other distance, either larger or smaller. ◀

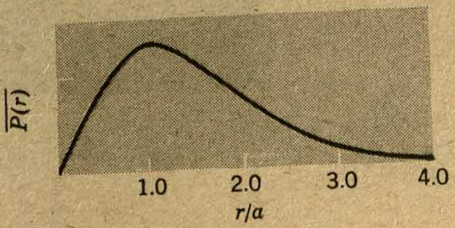


Fig. 48–9 The probability function for the ground state of the hydrogen atom, as calculated from wave mechanics. The separation between the nucleus and the electron is r; a is the radius of the first Bohr orbit (0.529×10^{-10} meter), used here merely as a convenient unit of distance.

48-5 The Uncertainty Principle

Only those quantities that can be measured have any real meaning in physics. If we could focus a "super" microscope on an electron in an atom and see it moving around in an orbit, we would declare that such orbits have meaning. However, we shall show that *it is fundamentally impossible to make such an observation*—even with the most ideal instruments that could conceivably be constructed. Therefore, we declare that such orbits have no physical meaning.

We observe the moon traveling around the earth by means of the sunlight that it reflects in our direction. Now light transfers linear momentum to an object from which it is reflected. In principle, this reflected light would disturb the course of the moon in its orbit, although a little thought shows that this disturbing effect is negligible.

For electrons the situation is quite different. Here, too, we can hope to "see" the electron only if we reflect light, or another particle, from it. In this case the recoil that the electron experiences when the light (photon) bounces from it completely alters the electron's motion in a way that cannot be avoided or even corrected for.

It is not surprising that the probability curve of Fig. 48-9 is the most detailed information that we can hope to obtain, by measurement, about the distribution of negative charge in the hydrogen atom. If orbits such as those envisaged by Bohr existed, they would be broken up completely in our attempts to verify their existence. Under these circumstances, we prefer to say that it is the probability function, and not the orbits, that represents physical reality.

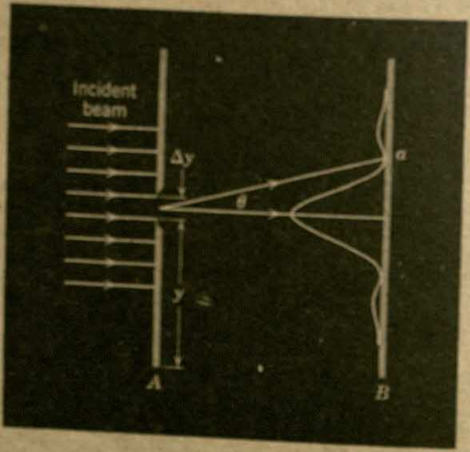


Fig. 48-10 An incident beam of electrons is diffracted at the slit in screen *A*, forming a typical diffraction pattern on screen *B*. If the slit is made narrower, the pattern becomes wider.

Our inherent inability to describe the motions of electrons in a classical way finds expression in the *uncertainty principle*, enunciated by Werner Heisenberg in 1927. To formulate this principle, consider a beam of monoenergetic electrons of speed v_0 moving from left to right in Fig. 48-10. Let us set ourselves the task of measuring the position of a particular electron in the vertical (y) direction and also its velocity component v_y in this direction. If we succeed in carrying out these measurements with unlimited accuracy, we can then claim to have established the position and motion of the electron (or one component of it at least) with precision. However, we shall see that *it is impossible to make these two measurements simultaneously with unlimited accuracy*.

To measure y we block the beam with an absorbing screen A in which we put a slit of width Δy. If an electron gets through the slit, its vertical position must be known to this accuracy. By making the slit narrower, we can improve the accuracy of this vertical position measurement as much as we wish.

Since the electron is a wave, it will undergo diffraction at the slit, and a photographic plate placed at B in Fig. 48-10 will reveal a typical *diffraction pattern*. The existence of this diffraction pattern means that there is an uncertainty in the values of v_y possessed by the electrons emerging from the slit. Let v_{ya} be the value of v_y that corresponds to an electron landing at the first minimum on the screen, marked by point a and described by a characteristic angle θ_a. We take v_{ya} as a *rough measure* of the uncertainty Δv_y in v_y for electrons emerging from the slit.

The first minimum in the diffraction pattern is given by Eq. 44-2, or

$$\sin\theta_a = \frac{\lambda}{\Delta y}.$$

If we assume that θ_a is small enough, we can write this equation as

$$\theta_a \cong \frac{\lambda}{\Delta y}. \tag{48-11}$$

To reach point a, v_{ya} $(= \Delta v_y)$ must be such that

$$\theta_a \cong \frac{\Delta v_y}{v_0}. \tag{48-12}$$

Combining Eqs. 48-11 and 48-12 leads to

$$\frac{\Delta v_y}{v_0} = \frac{\lambda}{\Delta y},$$

which we rewrite as

$$\Delta v_y\,\Delta y \cong \lambda v_0. \tag{48-13}$$

Now λ, the wavelength of the electron beam, is given by h/p or h/mv_0; substituting this into Eq. 48-13 yields

$$\Delta v_y\,\Delta y \cong \frac{hv_0}{mv_0}.$$

We rewrite this as

$$\Delta p_y\,\Delta y \cong h. \tag{48-14}$$

In Eq. 48-14 Δp_y $(= m\,\Delta v_y)$ is the uncertainty in our knowledge of the vertical momentum of the electrons; Δy is the uncertainty in our knowledge of their vertical position. The equation tells us that, since the product of these uncertainties is a constant, *we cannot measure p_y and y simultaneously with unlimited accuracy*.

If we want to improve our measurement of y (that is, if we want to reduce Δy), we use a finer slit. However (see Eq. 48-11), this will produce a wider diffraction pattern. A wider pattern means that our knowledge of the vertical momentum component of the electron has deteriorated, or, in other words, Δp_y has increased; this is exactly what Eq. 48-14 predicts.

The limits on measurement imposed by Eq. 48–14 have nothing to do with the crudity of our measuring instruments. We are permitted to postulate the existence of the finest conceivable measuring equipment. Equation 48–14 represents a *fundamental* limitation, imposed by nature.

Equation 48–14 is a derivation, for a special case, of a general principle known as the *uncertainty principle*. As applied to position and momentum measurements, it asserts that

$$\begin{aligned}\Delta p_x \, \Delta x &\gtrsim h\\ \Delta p_y \, \Delta y &\gtrsim h\\ \Delta p_z \, \Delta z &\gtrsim h. \end{aligned} \qquad (48\text{–}15)$$

Thus no component of the motion of an electron, free or bound, can be described with unlimited precision.

Planck's constant h probably appears nowhere that has more deep-seated significance than in Eq. 48–15. If this product had been zero instead of h, the classical ideas about particles and orbits would be correct; it would then be possible to measure both momentum and position with unlimited precision. The fact that h appears means that the classical ideas are wrong; the magnitude of h tells us under what circumstances these classical ideas must be replaced by quantum ideas. Gamow * has speculated, in an interesting and readable fantasy, what our world would be like if the constant h were much larger than it is, so that nonclassical ideas would be apparent to our sense perceptions.

▶ **Example 5.** An electron has a speed of 300 meters/sec, accurate to 0.01%. With what fundamental accuracy can we locate the position of this electron?

The electron momentum is

$$p = mv = (9.1 \times 10^{-31}\text{ kg})(300\text{ meters/sec}) = 2.7 \times 10^{-28}\text{ kg-m/sec}.$$

The uncertainty in momentum is given to be 0.01% of this, or

$$\Delta p = (0.0001)(2.7 \times 10^{-28}\text{ kg-m/sec}) = 2.7 \times 10^{-32}\text{ kg-m/sec}.$$

The minimum uncertainty in position, from Eq. 48–15, is

$$\begin{aligned}\Delta x = \frac{h}{\Delta p} &= \frac{6.6 \times 10^{-34}\text{ joule-sec}}{2.7 \times 10^{-32}\text{ kg-m/sec}}\\ &= 2.4\text{ cm}.\end{aligned}$$

If the electron momentum has really been determined by measurement to have the accuracy stated, there is no hope whatever that its position can be known to any better accuracy than that stated, namely about 1 in. The concept of the electron as a tiny dot is not very valid under these circumstances.

▶ **Example 6.** A bullet has a speed of 300 meters/sec, accurate to 0.01%. With what fundamental accuracy can we locate its position? Its mass is 50 gm (= 0.05 kg).

This example is the same as Example 5 in every respect save the mass of the particle involved. The momentum is

$$p = mv = (0.05\text{ kg})(300\text{ meters/sec}) = 15\text{ kg-m/sec}$$

and

$$\Delta p = (0.0001)(15\text{ kg-m/sec}) = 1.5 \times 10^{-3}\text{ kg-m/sec}.$$

Equation 48–15 yields

$$\Delta x = \frac{6.6 \times 10^{-34}\text{ joule-sec}}{1.5 \times 10^{-3}\text{ kg-m/sec}} = 4.4 \times 10^{-31}\text{ meter}.$$

---

* *Mr. Tompkins in Wonderland*, Macmillan, 1940.

This is so far beyond the possibility of measurement (a nucleus is only about 10^{-15} meter in diameter) that we can assert that for heavy objects like bullets the uncertainty principle sets no limits whatever on our measuring procedures. Once again the correspondence principle shows us how quantum physics reduces to classical physics under the appropriate circumstances. ◀

The uncertainty relation shows us why it is possible for both light and matter to have a dual, wave-particle, nature. It is because these two views, so obviously opposite to each other, can never be brought face to face in the same experimental situation. If we devise an experiment that forces the electron to reveal its wave character strongly, its particle character will always be inherently fuzzy. If we modify the experiment to bring out the particle character more strongly, the wave character necessarily becomes fuzzy. Matter and light are like coins that can be made to display either face at will but not both simultaneously. Niels Bohr first pointed out in his *principle of complementarity* how the ideas of wave and of particle complement rather than contradict each other.

QUESTIONS

1. How can the *wavelength* of an electron be given by $\lambda = h/p$ when the very presence of the momentum p in this formula implies that the electron is a particle?

2. How could Davisson and Germer be sure that the "54-volt" peak of Fig. 48–2 was a first-order diffraction peak, that is, that $m = 1$ in Eq. 48–2?

3. In a repetition of Thomson's experiment for measuring e/m for the electron (see Section 33–8), a beam of 10^4-ev electrons is collimated by passage through a slit of width 0.50 mm. Why is the beam-like character of the emerging electrons not destroyed by diffraction of the electron wave at this slit?

4. Why is the wave nature of matter not more apparent to our daily observations?

5. Apply the correspondence principle to the problem of a particle confined between rigid walls, showing that those features which seem "strange" (that is, the quantization of energy and the nonuniformity of the probability functions of Fig. 48–8) become undetectable experimentally at large quantum numbers.

6. In the $n = 1$ mode, for a particle confined between rigid walls, what is the probability that the particle will be found in a small volume element at the surface of either wall?

7. A standing wave can be viewed as the superposition of two traveling waves. Can you apply this to the problem of a particle confined between rigid walls, giving an interpretation in terms of the motion of the electron?

8. What is the physical significance of the wave function Ψ?

9. Why does the concept of Bohr orbits violate the uncertainty principle?

10. How can the predictions of wave mechanics be so exact if the only information we have about the positions of the electrons is statistical?

11. Make up some numerical examples to show the difficulty of getting the uncertainty principle to reveal itself during experiments with an object whose mass is about 1 gm.

12. Figure 48–8 shows that for $n = 3$ the probability function for a particle confined between rigid walls is zero at two points between the walls. How can the particle ever move across these positions? (Hint: Consider the implications of the uncertainty principle.)

13. The uncertainty principle can be stated in terms of angular quantities (compare Eq. 48–15) as

$$\Delta L \, \Delta\phi \gtrsim h$$

where ΔL is the uncertainty in the *angular* momentum and $\Delta\phi$ the uncertainty in the *angular* position. For electrons in atoms the angular momentum has definite quantized values, with no uncertainty whatever. What can we conclude about the uncertainty in the angular position and about the validity of the orbit concept?

PROBLEMS

1. A bullet of mass 40 gm travels at 1000 meters/sec. (*a*) What wavelength can we associate with it? (*b*) Why does the wave nature of the bullet not reveal itself through diffraction effects?

2. What wavelength do we associate with a beam of neutrons whose energy is 0.025 ev?

3. An electron and a photon each have a wavelength of 2.0 A. What are their (*a*) momenta and (*b*) energies?

4. The 50-Bev electron accelerator to be built at Stanford will provide an electron beam of small wavelength, suitable for probing the fine details of nuclear structure by scattering experiments. What will this wavelength be and how does it compare to the size of an average nucleus? (Hint: At these energies it is necessary to use the extreme relativistic relationship between momentum and energy, namely $p = E/c$. This is the same relationship used for light (Section 40–2) and is justified whenever the kinetic energy of a particle is very much greater than its rest energy m_0c^2, as in this case.)

5. In the experiment of Davisson and Germer (*a*) at what angles would the second- and third-order diffracted beams corresponding to the strong maximum of Fig. 48–2 occur, provided they are present, and (*b*) at what angle would the first-order diffracted beam occur if the accelerating potential were changed from 54 to 60 volts?

6. Make a plot of de Broglie wavelength against kinetic energy for (*a*) electrons and (*b*) protons. Restrict the range of energy values to those in which classical mechanics applies reasonably well. A convenient criterion is that the maximum kinetic energy on each plot be only about, say, 5% of the rest energy m_0c^2 for the particular particle.

7. What is the wavelength of a hydrogen atom moving with a velocity corresponding to the mean kinetic energy for thermal equilibrium at 20°C?

8. If an electron moves from a state represented by $n = 3$ in Fig. 48–8 to one represented by $n = 2$ in that figure, emitting electromagnetic radiation in the process, what are (*a*) the energy of the emitted single photon and (*b*) the corresponding wavelength?

9. A particle is confined between rigid walls separated by a distance l. What is the probability that it will be found within a distance $l/3$ of one wall (*a*) for $n = 1$, (*b*) for $n = 2$, (*c*) for $n = 3$, and (*d*) under the assumptions of classical physics?

10. In the ground state of the hydrogen atom show that the probability P_r that the electron lies *within* a sphere of radius r is given by

$$P_r = 1 - e^{-2r/a}\left(\frac{2r^2}{a^2} + \frac{2r}{a} + 1\right).$$

Does this yield expected values for (*a*) $r = 0$ and (*b*) $r = \infty$? (*c*) State clearly the difference in meaning between this expression and that given on p. 1117.

11. In the ground state of the hydrogen atom, what is the probability that the electron will lie within a sphere whose radius is that of the first Bohr orbit?

12. The uncertainty in the position of an electron is given as about 0.5 A, which is the radius of the first Bohr orbit in hydrogen. What is the uncertainty in the linear momentum of the electron?

13. Show that if the uncertainty in the location of a particle is equal to its de Broglie wavelength the uncertainty in its velocity is equal to its velocity.

14. The highest achievable resolving power of a microscope is limited only by the wavelength used; that is, the smallest detail that can be separated is about equal to the wavelength. Suppose one wishes to "see" inside an atom. Assuming the atom to have a diameter of 1.0 A, this means that we wish to resolve detail of separation about 0.1 A. (*a*) If an electron microscope is used, what minimum energy of electrons is needed? (*b*) If a light microscope is used, what minimum energy of photons is needed? (*c*) Which microscope seems more practical for this purpose? Explain.

15. A microscope using photons is employed to locate an electron in an atom to within a distance of 0.1 A. What is the uncertainty in the momentum of the electron located in this way?

The Differential Form of Maxwell's Equations and the Electromagnetic Wave Equation

SUPPLEMENTARY TOPIC V*

V–1 Introduction

In Chapter 39 we sought to make the existence of electromagnetic waves plausible by showing that such waves are consistent with Maxwell's equations as expressed in Table 38–3. Here we seek to start from Maxwell's equations and derive from them a differential equation whose solutions will describe electromagnetic waves. We will show directly that the speed c of such waves is given by Eq. 39–15, or $c = 1/\sqrt{\epsilon_0\mu_0}$.

We followed a similar program in Supplementary Topic III for mechanical waves on a stretched string. Starting from Newton's laws of motion we derived a differential equation (Eq. III–1) whose solutions (Eq. III–2) described such waves. We showed further that the speed v of these waves is given by Eq. III–3, or $v = \sqrt{F/\mu}$.

In Table 38–3 we wrote Maxwell's equations as

$$\epsilon_0 \oint \mathbf{E} \cdot d\mathbf{S} = q, \tag{V–1}$$

$$\oint \mathbf{B} \cdot d\mathbf{S} = 0, \tag{V–2}$$

$$\oint \mathbf{B} \cdot d\mathbf{l} = \mu_0(i + \epsilon_0\, d\Phi_E/dt), \tag{V–3}$$

and

$$\oint \mathbf{E} \cdot d\mathbf{l} = -d\Phi_B/dt. \tag{V–4}$$

---

* Supplementary Topics I to IV appear in Part I.

These equations are said to be written in integral form. The field variables **E** and **B**, which are usually the unknown quantities, appear in the integrands. Only in a few symmetric cases (see Sections 28–6 and 34–2, for example) can we "factor them out." In more general problems we cannot do so.

The situation is somewhat analogous to computing the density ρ of a body if we know its mass m and volume τ. In general these are related by the integral equation

$$m = \int \rho \, d\tau.$$

Only if ρ is a constant over all parts of the volume can we factor it out and write $\rho = m/\tau$.

To carry out our program it is desirable to recast Maxwell's equations in the form of equalities that apply at each *point* in space rather than as integrals that apply to various *regions* of space. In other words, we wish to convert Maxwell's equations from the integral form of Eqs. V–1 to 4 into *differential form*. We will then be able to relate **E** and **B** at a point to the charge density and current density at that point.

V–2 The Operator ∇

To transform Maxwell's equations into differential form we must deepen our understanding of vector methods and, in particular, become familiar with the vector operator ∇.

In Section 29–7 we saw how to obtain the components of the (vector) electrostatic field **E** at any point from the (scalar) potential function $V(x,y,z)$ by partial differentiation. Thus,

$$E_x = -\frac{\partial V}{\partial x}, \quad E_y = -\frac{\partial V}{\partial y}, \quad \text{and} \quad E_z = -\frac{\partial V}{\partial z}$$

so that the electrostatic field

$$\mathbf{E} = \mathbf{i}E_x + \mathbf{j}E_y + \mathbf{k}E_z$$

can be written as

$$\mathbf{E} = -\left(\mathbf{i}\frac{\partial V}{\partial x} + \mathbf{j}\frac{\partial V}{\partial y} + \mathbf{k}\frac{\partial V}{\partial z}\right). \tag{V–5}$$

We can write Eq. V–5 in compact vector notation as

$$\mathbf{E} = -\nabla V,$$

where ∇ ("del") is a vector operator defined as

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}. \tag{V–6}$$

This operator is useful in dealing with scalar and vector fields (see Sections 16–8 and 18–7 for examples of such fields). Given any scalar field ψ we may form a vector field, called the gradient of ψ and written as grad ψ or $\nabla\psi$, simply by applying the operator ∇ to ψ. Given a vector field $\mathbf{U} = U_x\mathbf{i} + U_y\mathbf{j} + U_z\mathbf{k}$ we may apply the operator ∇ to it in two different ways. One way is to take the dot product of ∇ and **U**, yielding the scalar field called the divergence of **U** and written as div **U** or $\nabla \cdot \mathbf{U}$. The other way is to take the cross product of ∇ and **U**, yielding the vector field called the curl of **U** and written curl **U** or $\nabla \times \mathbf{U}$. These operations may be summarized as

$$\operatorname{grad} \psi \equiv \nabla \psi = \mathbf{i}\frac{\partial \psi}{\partial x} + \mathbf{j}\frac{\partial \psi}{\partial y} + \mathbf{k}\frac{\partial \psi}{\partial z}$$

$$\operatorname{div} \mathbf{U} \equiv \nabla \cdot \mathbf{U} = \frac{\partial U_x}{\partial x} + \frac{\partial U_y}{\partial y} + \frac{\partial U_z}{\partial z}$$

$$\begin{aligned}\operatorname{curl} \mathbf{U} \equiv \nabla \times \mathbf{U} = \mathbf{i}\left(\frac{\partial U_z}{\partial y} - \frac{\partial U_y}{\partial z}\right) \\ + \mathbf{j}\left(\frac{\partial U_x}{\partial z} - \frac{\partial U_z}{\partial x}\right) \\ + \mathbf{k}\left(\frac{\partial U_y}{\partial x} - \frac{\partial U_x}{\partial y}\right)\end{aligned}$$

Note that grad ψ and curl **U** are vectors, whereas div **U** is a scalar. The student can gain some familiarity with these operations by the following exercises: (1) prove that curl (grad ψ) = 0 and (2) prove that div (curl **U**) = 0.

Another frequently occuring operator is ∇^2 ("del squared"). It is simply $\nabla \cdot \nabla$, or, as the student can show from Eq. V–6,

$$\nabla^2 = \nabla \cdot \nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

When we apply ∇^2 to a scalar field ψ, we obtain

$$\nabla^2 \psi = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} \tag{V–10}$$

For a vector field **U**, the operation $\nabla^2 \mathbf{U}$ is defined as

$$\nabla^2 \mathbf{U} = \mathbf{i}\frac{\partial^2 U_x}{\partial x^2} + \mathbf{j}\frac{\partial^2 U_y}{\partial y^2} + \mathbf{k}\frac{\partial^2 U_z}{\partial z^2}. \tag{V–11}$$

As an exercise the student can show that curl (curl **U**) = $-\nabla^2 \mathbf{U}$ + grad (div **U**).

V–3 Maxwell's Equations in Differential Form—I

In this section we show how to cast the first two of Maxwell's equations (Eqs. V–1, 2) into differential form. Let us apply Eq. V–1 to a differential volume element shaped like a rectangular parallelepiped and containing a point P at (and near) which an electric field exists (see Fig. V–1*a*). Point P is located at x, y, z in the reference frame of Fig. V–1*b* and the edges of the parallelepiped have lengths dx, dy, and dz.

We can write the surface area vector for the rear face of the parallelepiped as $d\mathbf{S} = -\mathbf{i}\,dy\,dz$. The minus sign enters because $d\mathbf{S}$ is defined to point in the direction of the *outward* normal, which is defined by $-\mathbf{i}$. For the front face we have $d\mathbf{S} = +\mathbf{i}\,dy\,dz$.

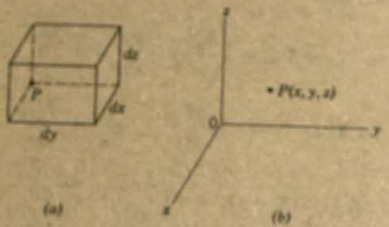


Fig. V–1

If the electric field at the rear face is **E**, that at the front face, which is a distance dx away from the rear face, is $\mathbf{E} + (\partial \mathbf{E}/\partial x)\, dx$, the latter term representing the change in **E** associated with the change dx in x.

The flux through the entire surface of the parallelepiped is $\oint \mathbf{E} \cdot d\mathbf{S}$ and the contribution to this flux due to these two faces alone is

$$(\mathbf{E}) \cdot (-\mathbf{i}\, dy\, dz) + \left(\mathbf{E} + \frac{\partial \mathbf{E}}{\partial x}\, dx\right) \cdot (+\mathbf{i}\, dy\, dz)$$

$$= dx\, dy\, dz \left(\frac{\partial \mathbf{E}}{\partial x} \cdot \mathbf{i}\right) = dx\, dy\, dz\, \frac{\partial}{\partial x} (\mathbf{E} \cdot \mathbf{i})$$

$$= dx\, dy\, dz\, \frac{\partial E_x}{\partial x}.$$

With similar contributions from the other four faces the total electric flux becomes

$$\oint \mathbf{E} \cdot d\mathbf{S} = dx\, dy\, dz \left(\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z}\right).$$

From Eq. V–8 we may write this as

$$\oint \mathbf{E} \cdot d\mathbf{S} = dx\, dy\, dz\, \operatorname{div} \mathbf{E}. \tag{V–12}$$

Now the right-hand side of Eq. V–1, which gives the charge enclosed by the surface, may be written in general as $q = \int \rho\, d\tau$ and, in particular, for the differential volume element at P, as

$$q = \rho\, dx\, dy\, dz, \tag{V–13}$$

where ρ is the charge per unit volume at P. Substituting Eqs. V–12 and 13 into Eq. V–1 and canceling the common factor $dx\, dy\, dz$, we have finally

$$\epsilon_0 \operatorname{div} \mathbf{E} = \rho, \tag{V–14}$$

which is Maxwell's first equation (Eq. V–1) in differential form.

Using the same technique, we can express Maxwell's second equation (Eq. V–2) in differential form as

$$\operatorname{div} \mathbf{B} = 0. \tag{V–15}$$

V-4 Maxwell's Equations in Differential Form—II

We now seek to transform Maxwell's third and fourth equations (Eqs. V-3, 4) into differential form. We start by applying Eq. V-3 to a differential surface element of rectangular shape at a point P in some region of a magnetic field, as shown in Fig. V-2*a*. The point P is located at x, y, z in the reference frame of Fig. V-2*b* and the sides of the rectangle, which is parallel to the x-y plane, have lengths dx and dy. Going around the path, as shown by the arrows, we have

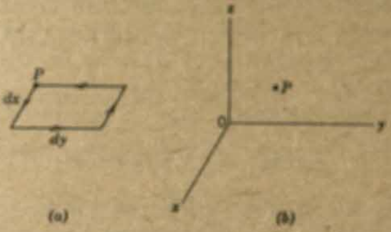


Fig. V-2

$$
\begin{aligned}
\oint \mathbf{B}\cdot d\mathbf{l} &= \mathbf{B}\cdot(-\mathbf{j}\,dy) && \text{(rear side)}\\
&+ \mathbf{B}\cdot(+\mathbf{i}\,dx) && \text{(left side)}\\
&+ \left(\mathbf{B} + \frac{\partial \mathbf{B}}{\partial x}\,dx\right)\cdot(+\mathbf{j}\,dy) && \text{(front side)}\\
&+ \left(\mathbf{B} + \frac{\partial \mathbf{B}}{\partial y}\,dy\right)\cdot(-\mathbf{i}\,dx), && \text{(right side)}
\end{aligned}
$$

where $\mathbf{B}$ is the magnetic induction at P.

Collecting terms we obtain

$$
\begin{aligned}
\oint \mathbf{B}\cdot d\mathbf{l} &= dx\,dy\left(\frac{\partial \mathbf{B}}{\partial x}\cdot\mathbf{j} - \frac{\partial \mathbf{B}}{\partial y}\cdot\mathbf{i}\right)\\
&= dx\,dy\left[\frac{\partial}{\partial x}(\mathbf{B}\cdot\mathbf{j}) - \frac{\partial}{\partial y}(\mathbf{B}\cdot\mathbf{i})\right]\\
&= dx\,dy\left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}\right). \qquad \text{(V-16)}
\end{aligned}
$$

Now in the right-hand side of Eq. V-3 i is the current enclosed by the path and $d\Phi_E/dt$ is the change in electric flux through the enclosed surface. Hence, if $\mathbf{J}$ is taken to represent the current density and $d\mathbf{S}(=\mathbf{k}\,dx\,dy)$ the surface area vector, we can write

$$i = \mathbf{J}\cdot d\mathbf{S} = \mathbf{J}\cdot(\mathbf{k}\,dx\,dy) = dx\,dy\,J_z \qquad \text{(V-17)}$$

and

$$\frac{d\Phi_E}{dt} = \frac{\partial \mathbf{E}}{\partial t}\cdot d\mathbf{S} = \frac{\partial \mathbf{E}}{\partial t}\cdot(\mathbf{k}\,dx\,dy)$$

or

$$\frac{d\Phi_E}{dt} = \frac{\partial E_z}{\partial t}\,dx\,dy. \qquad \text{(V-18)}$$

Substituting Eqs. V–16, 17, and 18 into Eq. V–3 and canceling the common factor $dx\,dy$, we get

$$\frac{1}{\mu_0}\left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}\right) = J_z + \epsilon_0 \frac{\partial E_z}{\partial t} \tag{V–19}$$

We could have proceeded exactly as above if we had started with a rectangle parallel to the y-z plane or one parallel to the z-x plane. Each rectangle would have given us a different component of an arbitrarily oriented differential surface at P. Equation V–19 is obviously the z-component equation corresponding to Eq. V–3. If we multiply it by **k** and add to it the two similar vector equations, which may be obtained by cyclically permuting x, y, z and **i**, **j**, **k**, corresponding to the x-component and y-component equations, we obtain

$$\text{curl}\,\mathbf{B} = \mu_0(\mathbf{J} + \epsilon_0\,\partial\mathbf{E}/\partial t), \tag{V–20}$$

which is the third Maxwell equation in differential form.

Similarly, starting with Eq. V–4, we may show that

$$\text{curl}\,\mathbf{E} = -\partial\mathbf{B}/\partial t, \tag{V–21}$$

which is the fourth Maxwell equation in differential form.

We have derived four differential equations (see V-22 to V-25 below) from the four integral equations (V-1 to V-4). It can be shown that the integral equations can be derived from the differential equations, that is, the two sets of equations are *equivalent*

V–5 The Wave Equation

We have now obtained from their integral form the four basic equations of electromagnetism, Maxwell's equations, in differential form. Corresponding to the integral equations, Eqs. V–1, 2, 3, and 4 respectively, we have

$$\epsilon_0\,\text{div}\,\mathbf{E} = \rho, \tag{V–22}$$

$$\text{div}\,\mathbf{B} = 0, \tag{V–23}$$

$$\text{curl}\,\mathbf{B} = \mu_0(\mathbf{J} + \epsilon_0\,\partial\mathbf{E}/\partial t), \tag{V–24}$$

$$\text{curl}\,\mathbf{E} = -\partial\mathbf{B}/\partial t, \tag{V–25}$$

which are four coupled partial differential equations. They apply at each point of space in an electromagnetic field.

We will now derive the wave equation for electromagnetic waves in free space. In free space, the charge density ρ and the current density **J** are zero, so that the Maxwell equations there become

$$\text{div}\,\mathbf{E} = 0,$$

$$\text{div}\,\mathbf{B} = 0,$$

$$\text{curl}\,\mathbf{E} = -\,\partial\mathbf{B}/\partial t,$$

and

$$\text{curl}\,\mathbf{B} = \mu_0\epsilon_0\,\partial\mathbf{E}/\partial t.$$

Let us take the curl of the equation for curl **E**; we obtain

$$\text{curl curl}\,\mathbf{E} = -\,\text{curl}\,\frac{\partial\mathbf{B}}{\partial t} = -\frac{\partial}{\partial t}\,\text{curl}\,\mathbf{B}.$$

But, from above, curl $\mathbf{B} = \mu_0\epsilon_0(\partial\mathbf{E}/\partial t)$, so that

$$\text{curl curl } \mathbf{E} = -\mu_0\epsilon_0 \frac{\partial^2\mathbf{E}}{\partial t^2} \tag{V-26}$$

From the exercise in V–2, we know that curl curl $\mathbf{E} = -\nabla^2 E +$ grad div $\mathbf{E}$ and from above that div $\mathbf{E} = 0$. Thus,

$$\text{curl curl } \mathbf{E} = -\nabla^2\mathbf{E}. \tag{V-27}$$

Combining Eqs. V–26 and V–27, we obtain finally

$$\nabla^2\mathbf{E} = \mu_0\epsilon_0 \frac{\partial^2\mathbf{E}}{\partial t^2}. \tag{V-28a}$$

The student, proceeding as above, should be able to show that **B** satisfies the same equation, or

$$\nabla^2\mathbf{B} = \mu_0\epsilon_0 \frac{\partial^2\mathbf{B}}{\partial t^2}. \tag{V-28b}$$

Equations V–28 are the equations of electromagnetic wave motion. Being vector equations, they are equivalent to six scalar equations, one for each component of **E** and of **B**.

There are many solutions of Eqs. V–28, corresponding to different kinds of electromagnetic waves—plane, spherical, and cylindrical waves being three examples. Let us consider a solution in which two components of **E** and two of **B** vanish, that is, in which

$$E_x = E_z = 0 \qquad \text{and} \quad B_x = B_y = 0.$$

Equations V–28 are satsified for these assumptions. For the nonvanishing components, E_y and B_z, Eqs. V–28 reduce to (see Eq. V–11)

$$\frac{\partial^2 E_y}{\partial x^2} + \frac{\partial^2 E_y}{\partial y^2} + \frac{\partial^2 E_y}{\partial z^2} = \mu_0\epsilon_0 \frac{\partial^2 E_y}{\partial t^2} \tag{V-29a}$$

and

$$\frac{\partial^2 B_z}{\partial x^2} + \frac{\partial^2 B_z}{\partial y^2} + \frac{\partial^2 B_z}{\partial z^2} = \mu_0\epsilon_0 \frac{\partial^2 B_z}{\partial t^2} \tag{V-29b}$$

If we make the additional assumption that E_y and B_z are functions of x and t only, the simplified wave equation that results, namely

$$\frac{\partial^2 E_y}{\partial x^2} = \mu_0\epsilon_0 \frac{\partial^2 E_y}{\partial t^2} \quad \text{and} \quad \frac{\partial^2 B_z}{\partial x^2} = \mu_0\epsilon_0 \frac{\partial^2 B_z}{\partial t^2}, \tag{V-29c}$$

is similar to Eq. III-1 for the vibrating string.

A solution to these equations, as the student may verify by substitution, is

$$E_y = E_m \sin(kx - \omega t) \tag{V-30a}$$

and

$$B_z = B_m \sin(kx - \omega t). \tag{V-30b}$$

We interpret Eqs. V–30 as an electromagnetic wave traveling in the positive x-direction, as in Fig. 39–11, with a speed $c = \omega/k$. The student can show that substituting Eq. V–30a into Eq. V–29a (or Eq. V–30b into Eq. V–29b) yields

$$c = \omega/k = 1/\sqrt{\epsilon_0\mu_0},$$

which (see Eq. 39–15) gives the speed of electromagnetic waves in free space.

Supplementary Problems

Chapter 26

1. Two free point charges $+q$ and $+4q$ are a distance l apart. A third charge is so placed that the entire system is in equilibrium. Find the location, magnitude, and sign of the third charge.

2. If the balls of Fig. 26-7 are conducting, what happens to them after one is discharged? Find the new equilibrium separation.

3. Two identical conducting spheres, having charges of opposite sign, attract each other with a force of 0.108 nt when separated by 0.5 meter. The spheres are connected by a conducting wire, which is then removed, and thereafter repel each other with a force of 0.036 nt. What were the initial charges on the spheres?

4. A particle of charge $-q$ and mass m moves in a circular orbit about a fixed charge $+Q$. (*a*) Show that the "distance cubed $\propto$ period squared" law,

$$r^3 = \frac{Qq}{16\pi^3\epsilon_0 m} T^2$$

is satisfied. Note that the proportionality constant depends on the property (q/m) of the orbiting particle. (*b*) What is the corresponding situation when the force is gravitational rather than electrical?

5. An electron is projected with an initial speed of 3.24×10^5 meter/sec directly toward a proton which is essentially at rest. If the electron is initially a very great distance from the proton, at what distance from the proton is its speed instantaneously equal to twice its initial value? (*Hint:* Use the work-energy theorem.)

6. A "dipole" is formed from a rod of length $2a$ and two charges, $+q$ and $-q$. Two such dipoles are oriented as shown in Fig. 26-9, their centers being separated by the

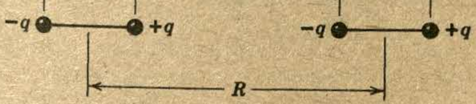


Fig. 26-9

distance R. (*a*) Calculate the force exerted on the left dipole. (*b*) For $R \gg a$, show that the magnitude of the force exerted on the left dipole is approximately given by

$$F = \frac{3p^2}{2\pi\epsilon_0 R^4}$$

where $p = 2qa$ is the "dipole moment."

Chapter 27

1. A uniform vertical field **E** is established in the space between two large parallel plates. In this field one suspends a small conducting sphere of mass m from a string of length l. Find the period of this pendulum when the sphere is given a charge $+q$ if the lower plate is charged positively; is charged negatively.

2. A charge $q = 3.0 \times 10^{-6}$ coul is 30 cm from a small dipole along its perpendicular bisector. The magnitude of the force on the charge is 5.0×10^{-6} nt. Show on a diagram (*a*) the direction of the force on the charge, (*b*) the direction of the force on the dipole, and (*c*) determine the magnitude of the force on the dipole.

3. A thin glass rod is bent into a semicircle of radius R. A charge $+Q$ is uniformly distributed along the upper half and a charge $-Q$ is uniformly distributed along the lower half, as shown in Fig. 27-25. Find the electric field **E** at P, the center of the semicircle.

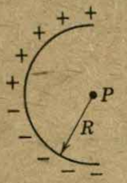


Fig. 27-25

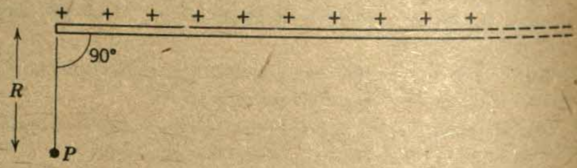


Fig. 27-26

4. A "semi-infinite" insulating rod (Fig. 27-26) carries a constant charge per unit length of λ. Show that the electric field at the point P makes an angle of 45° with the rod. This result is independent of the distance R.

5. Find the frequency of oscillation of an electric dipole, of moment p and rotational inertia I, for small amplitudes of oscillation about its equilibrium position in a uniform electric field of strength E.

6. An electric dipole of moment **p** is placed parallel to an electric field line along the y-axis in a nonuniform electric field (Fig. 27-27). The magnitude of the field **E** varies uniformly along the y-direction as shown. (*a*) Show that the magnitude of the force on the dipole is $p\,dE/dy$. (*b*) What is the direction of the force?

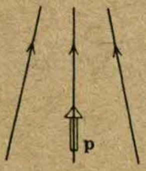


Fig. 27-27

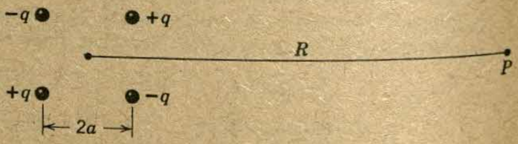


Fig. 27-28

7. One type of "electric quadrupole" is formed by four charges located at the vertices of a square of side $2a$. Point P lies a distance R from the center of the quadrupole on a line parallel to two of sides of the square as shown in Fig. 27-28. For $R \gg a$, show that

the electric field at P is approximately given by

$$E = \frac{3(2qa^2)}{4\pi\epsilon_0 R^4}.$$

(*Hint:* Treat the quadrupole as two dipoles.)

Chapter 28

1. It is found experimentally that the electric field in a large region of the earth's atmosphere is directed vertically down. At an altitude of 300 meters the field is 60 volts/meter and at an altitude of 200 meters it is 100 volts/meter. Find the net amount of charge contained in a cube 100 meters on edge located between 200 and 300 meters altitude. Neglect the curvature of the earth.

2. Suppose that an electric field in some region is found to have a constant direction but to be decreasing in strength in that direction. What do you conclude about the charge in the region?

3. Two concentric conducting spherical shells have radii $R_1 = 0.145$ meter and $R_2 = 0.207$ meter. The inner sphere bears a charge -6.00×10^{-2} coul. An electron escapes from the inner sphere with negligible speed. Assuming that the region between the spheres is a vacuum, compute the speed with which the electron strikes the outer sphere.

4. (*a*) Two identical nonconducting spheres have radius r and are fixed with their centers a distance $R > 2r$ apart. If each sphere has a total charge q *uniformly distributed* on its surface, what is the magnitude of the electric force that either sphere exerts on the other? (*b*) Suppose instead that the spheres are conductors, the same total charge on each still being q. Will the electric force that either sphere exerts on the other in this case be greater than, less than, or equal to the force in case (*a*)? Explain.

5. The spherical region $a < r < b$ carries a charge per unit volume of $\rho = A/r$, where A is constant. At the center ($r = 0$) of the enclosed cavity is a point charge Q. What should the value of A be so that the electric field in the region $a < r < b$ has constant magnitude?

6. A solid insulating sphere carries a uniform charge per unit volume of ρ. Let $\mathbf{r}$ be the vector from the center of the sphere to a general point P within the sphere. (*a*) Show that the electric field at P is given by $\mathbf{E} = \rho\mathbf{r}/3\epsilon_0$. (*b*) A spherical "cavity" is removed from the above sphere, as shown in Fig. 28–25. Using superposition concepts show that the electric field at all points within the cavity is $\mathbf{E} = \rho\mathbf{a}/3\epsilon_0$ (uniform field), where $\mathbf{a}$ is the vector connecting the center of the sphere with the center of the cavity. Note that both these results are independent of the radii of the sphere and the cavity.

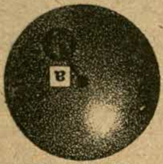


Fig. 28–25

Chapter 29

1. What is the charge density on the surface of a conducting sphere of radius 0.15 meter whose potential is 200 volts?

2. Two identical conducting spheres of radius $r = 0.15$ meter are separated by a distance $a = 10.0$ meters. What is the charge on each sphere if the potential of one is +1500 volts and if the other is −1500 volts?

3. Two conducting spheres, one of radius 6.0 cm and the other of radius 12.0 cm, each have a charge of 3×10^{-8} coul and are very far apart. If the spheres are connected by a conducting wire, find (*a*) the direction of motion and the magnitude of the charge transferred and (*b*) the final charge on and potential of each sphere.

4. In the rectangle shown in Fig. 29–29, the sides have lengths 5.0 cm and 15.0 cm, $q_1 = -5.0 \times 10^{-6}$ coul and $q_2 = +2.0 \times 10^{-6}$ coul. (*a*) What is the electric potential at corner *B*? At corner *A*? (*b*) How much work is involved in moving a third charge $q_3 = +3.0 \times 10^{-6}$ coul *from B to A* along a diagonal of the rectangle? (*c*) In this process, is external work converted into electrostatic potential energy or vice versa? Explain.

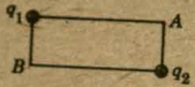


Fig. 29–29

5. Show that the potential energy of an assembly of four charges, each of charge q, in the configuration of a triangular pyramid, a_0 on an edge, is $6 \times (q^2/a_0)/4\pi\epsilon_0$.

6. Three charges of +0.1 coul each are placed on the corner of an equilateral triangle, 1.0 meter on a side. If energy is supplied at the rate of 1.0 kw, how many days would be required to move one of the charges onto the midpoint of the line joining the other two?

7. Show, using the fact that an electrostatic field is a conservative field, that one cannot create an electric field in which all the lines of force are straight parallel lines whose density (number per unit cross-sectional area) changes in a direction at right angles to the lines of force.

8. Two line charges are parallel to the z-axis; one, of charge per unit length $+\lambda$, is a distance a to the right of this axis, and the other, of charge per unit length $-\lambda$, is a distance a to the left of this axis (the lines and z-axis being in the same plane). Sketch some of the surfaces of equipotential.

9. A particle of (positive) charge Q is assumed to have a fixed position at P. A second particle of mass m and (negative) charge $-q$ moves at constant speed in a circle of radius r_1, centered at P. Derive an expression for the work W that must be done by an external agent on the second particle in order to increase the radius of the circle of motion, centered at P, to r_2. Express W in terms of quantities chosen from among m, r_1, r_2, q, Q, and ϵ_0 only.

10. A particle of charge Q is kept in a fixed position at a point P and a second particle of mass m, having the same charge Q, is initially held at rest a distance r_1 from P. The second particle is then released and is repelled from the first one. Determine its velocity at the instant it is a distance r_2 from P. Let $Q = 3.1 \times 10^{-6}$ coul, $m = 2.0 \times 10^{-5}$ kg, $r_1 = 9.0 \times 10^{-4}$ meter, and $r_2 = 25 \times 10^{-4}$ meter.

11. A particle of mass m, charge $q > 0$, and initial kinetic energy K is projected (from "infinity") toward a heavy nucleus of charge Q, assumed to have a fixed position

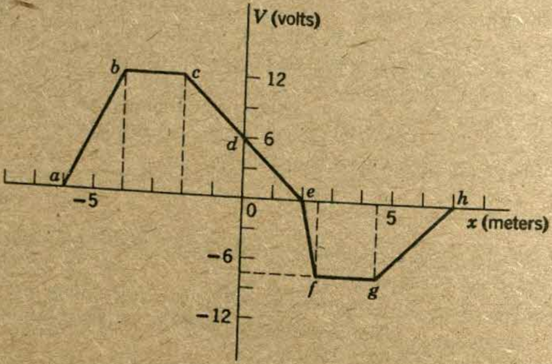


Fig. 29–30

in our reference frame. (*a*) If the "aim is perfect," how close to the center of the nucleus is the particle when it comes instantaneously to rest? (*b*) With a particular imperfect aim the particle's closest approach to the nucleus is *twice* the distance determined in part (*a*). Determine the *speed* of the particle at this closest distance of approach.

12. The electric potential varies along the x-axis as shown in the graph of Fig. 29–30. For each of the intervals shown (ignore the behavior at the end points of the intervals) determine the x-component of the electric field and plot E_x vs. x.

13. A charge per unit length λ is distributed uniformly along a straight-line segment of length L. (*a*) Determine the electrostatic potential (chosen to be zero at infinity) at a point P a distance y from one end of the charged segment and in line with it (see Fig. 29–31). (*b*) Use the result of (*a*) to compute the component of the electric field intensity at P in the y-direction (along the line). (*c*) Determine the component of the electric field intensity at P in a direction perpendicular to the straight line.

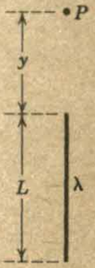


Fig. 29–31

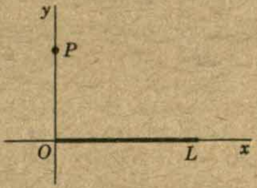


Fig. 29–32

14. On a thin rod of length L lying along the x-axis with one end at the origin ($x = 0$), as in Fig. 29–32, there is distributed a charge per unit length given by $\lambda = kx$, where k is a constant. (*a*) Taking the electrostatic potential at infinity to be zero, find V at the point P on the y-axis. (*b*) Determine the vertical component, E_y, of the electric field intensity at P from the result of part (*a*) and also by direct calculation. (*c*) Why cannot E_x, the horizontal component of the electric field at P, be found using the result of part (*a*)?

Chapter 30

1. Show that the capacitance of two oppositely charged metal spheres of the same radius, when far apart, is one-half the capacitance of one isolated sphere.

2. Two metallic spheres, radii a and b, are connected by a thin wire. Their separation is large compared with their dimensions. A charge Q is put onto this system. (*a*) How much charge resides on each sphere? (*b*) Apply the definition of capacitance to show that the capacitance of this system is $C = 4\pi\epsilon_0 (a + b)$.

3. N identical spherical drops of liquid are charged to the same potential V. One large drop is formed by combining these. Show that the potential of the large drop is $N^{2/3}V$.

4. Charges q_1, q_2, q_3 are placed on capacitors of capacitance C_1, C_2, C_3 respectively, arranged in series as shown in Fig. 30–28. Switch S is then closed. What are the final charges q_1', q_2', q_3' on the capacitors?

5. When switch S is thrown to the left in Fig. 30–29, the plates of the capacitor of capacitance C_1 acquire a potential difference V_0. C_2 and C_3 are initially uncharged.

The switch is now thrown to the right. What are the final charges q_1, q_2, q_3 on the corresponding capacitors?

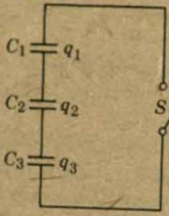


Fig. 30–28

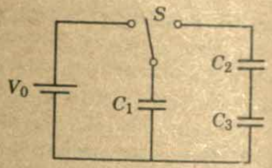


Fig. 30–29

6. A geiger tube is made of two long concentric metal cylinders with a gas of dielectric constant κ between them. Neglecting end effects, use Gauss' law to calculate the capacitance of this configuration. The center rod has a radius a, the surrounding tube a radius b, and the length $l \gg b$.

7. A parallel plate capacitor has plates of area 0.12 m^2 and a separation of 1.2 cm. A battery charges the plates to a potential difference of 120 volts and is then disconnected. A dielectric slab of thickness 0.4 cm and dielectric constant 4.8 is then placed symmetrically between the plates. In terms of ϵ_0 (*a*) find the capacitance before the slab is inserted. (*b*) what is the capacitance with the slab in place? (*c*) what is the free charge q before and after the slab is inserted? (*d*) determine the electric field strength in the space between the plates and dielectric. (*e*) what is the electric field strength in the dielectric? (*f*) with the slab in place what is the potential difference across the plates? (*g*) how much external work is involved in the process of inserting the slab?

Chapter 31

1. A steady beam of alpha particles ($q = 2e$) traveling with constant kinetic energy 20 MeV carries a current 0.25×10^{-6} ampere. (*a*) If the beam is directed perpendicular to a plane surface, how many alpha particles strike the surface in 3.0 sec? (*b*) At any instant, how many alpha particles are there in a given 20-cm length of the beam? (*c*) Through what potential difference was it necessary to accelerate each alpha particle from rest to bring it to an energy of 20 MeV?

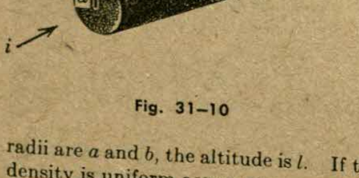


Fig. 31–10

2. List similarities and differences between the flow of charge and the flow of a fluid; between the flow of charge and the conduction of heat.

3. Explain why the momentum which conduction electrons transfer to the ions in a metal conductor does not give rise to a resultant force on the conductor.

4. A resistor is in the shape of a truncated right circular cone (Fig. 31–10). The end radii are a and b, the altitude is l. If the taper is small, we may assume that the current density is uniform across any cross section. (*a*) Calculate the resistance of this object. (*b*) Show that your answer reduces to $\rho(l/A)$ for the special case of zero taper ($a = b$).

5. Conductors A and B, having equal lengths of 40 meters and a cross-sectional area of 0.10 m^2, are connected in series. A potential of 60 volts is applied across the terminal points of the connected wires. The resistances of the wires are 40 and 20 ohms respectively. Determine: (*a*) the resistivities of the two wires; (*b*) the magnitude of the elec-

tric field in each wire; (*c*) the current density in each wire; (*d*) the potential difference applied to each conductor.

6. A 1250-watt radiant heater is constructed to operate at 115 volts. (*a*) What will be the current in the heater? (*b*) What is the resistance of the heating coil? (*c*) How many kilocalories are created in our hour by the heater?

7. An iron wire (diameter 1 mm, length 10 cm) is placed in an evacuated chamber. Estimate the equilibrium temperature of the wire if it carries a current of 10 amp. Assume that all heat transfer is by radiation and that the surface of the wire radiates according to Eq. 47–2. Take the temperature of the chamber walls to be 27° C. List any additional assumptions that you use.

Chapter 32

1. Two batteries having the same emf $\mathcal{E}$ but different internal resistances r_1 and r_2 are connected in series to an external resistance R. Find the value of R that makes the potential difference zero between the terminals of the first battery.

2. The section of circuit AB (see Fig. 32–26) absorbs power $P = 50.0$ watts and a current $i = 1.0$ amp passes through it in the indicated direction. (*a*) What is the potential difference between A and B? (*b*) If the element C does not have internal resistance, what is its emf? (*c*) What is its polarity?

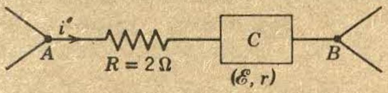


Fig. 32–26

3. Two light bulbs, one of resistance R and the other of resistance $r (< R)$, are connected (*a*) in parallel and (*b*) in series. Which bulb is brighter?

4. N identical batteries of emf $\mathcal{E}$ and internal resistance r may be connected all in series or all in parallel. Show that each arrangement will give the same current in an external resistor R if $R = r$.

5. Twelve resistors, each of resistance R ohms, form a cube (see Fig. 32–27). (*a*) Find R_{AB}, the resistance of an edge. (*b*) Find R_{BC}, the equivalent resistance of a face. (*c*) Find R_{AC}, the equivalent resistance of the body diagonal.

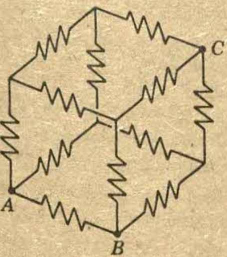


Fig. 32–27

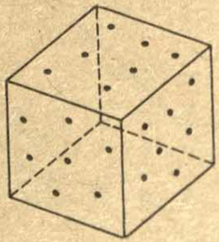
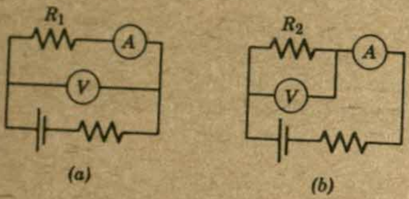


Fig. 32–28

6. What internal connections (resistors only) can account for the fact that there are 2 ohms resistance between any pair of the terminals on a box with N terminals? (See Fig. 32–28.)

7. A voltmeter and an ammeter are used to determine two unknown resistances R_1 and R_2, one determination by each of the two methods shown in Fig. 32–29. The voltmeter resistance is 307 ohms and the ammeter resistance is 3.62 ohms; in method (*a*) the ammeter reads 0.317 amp and the voltmeter reads 28.1 volts, whereas in method (*b*) the ammeter reads 0.356 amp and the voltmeter 23.7 volts. Compute R_1 and R_2.



(*a*) (*b*)

Fig. 32–29

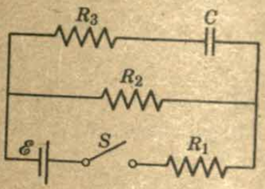


Fig. 32–30

8. An *RC* circuit is discharged by closing a switch at time $t = 0$. The initial potential difference across the capacitor is 100 volts. If the potential difference has decreased to 1.0 volt after 10 sec, (*a*) what will the potential difference be 20 sec after $t = 0$? (*b*) What is the time constant of the circuit?

9. A capacitor with capacitance $C = 1.0\ \mu f$ and initial stored energy $U_0 = 0.5$ joule is discharged through a resistance $R = 1.0 \times 10^6$ ohms. (*a*) What is the initial charge on the capacitor? (*b*) What is the current through the resistor when the discharge starts? (*c*) Determine V_c, the voltage across the capacitor, and V_R, the voltage across the resistor, as a function of time. (*d*) Express the rate of joule heating in the resistor as a function of time.

10. In the circuit of Fig. 32–30, $\mathcal{E} = 1200$ volts, $C = 6.50\ \mu f$, $R_1 = R_2 = R_3 = 7.30 \times 10^5$ ohms. With C completely uncharged, the switch S is suddenly closed ($t = 0$). (*a*) Determine, for $t = 0$ and $t = \infty$, the currents through each resistor. (*b*) Draw qualitatively a graph of the potential drop V_2 across R_2 from $t = 0$ to $t = \infty$. (*c*) What are the numerical values of V_2 at $t = 0$ and $t = \infty$? (*d*) Give the physical meaning of "$t = \infty$" and state a rough, but significant, numerical lower bound, in seconds, for "$t = \infty$" in this case.

Chapter 33

1. Particles 1, 2, and 3 follow the paths shown in Fig. 33–25 as they pass through the magnetic field there. What can one conclude about each particle?

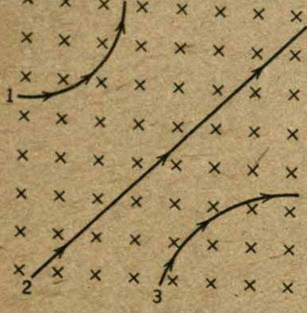


Fig. 33–25

2. (*a*) What speed would a proton need to circle the earth at the equator, if the earth's magnetic field is everywhere horizontal there and directed along longitudinal lines. Take the magnitude of the earth's magnetic field to be 0.41×10^{-4} weber/meter^2 at the equator. (*b*) Draw the velocity and magnetic induction vectors corresponding to this situation.

3. An electron is accelerated through a potential difference of 1000 volts and directed into a region between two parallel plates separated by 0.02 meter with a voltage difference of 100 volts between them. If the electron enters moving perpendicular to the electric field between the plates, what magnetic field is necessary perpendicular to both the electron path and the electric field so that the electron travels in a straight line?

4. A positive point charge Q travels in a straight line with constant speed through an evacuated region in which there is a uniform electric field **E** and a uniform magnetic field **B**. (*a*) If **E** is directed vertically up and the charge travels horizontally from north to south with speed v, determine the least value of the magnitude of **B** and the corresponding direction of **B**. (*b*) Explain why **B** is not uniquely determined when **E** and **v** alone are given. (*c*) Suppose the charge is a proton which enters the region after having been accelerated through a potential difference of 3.10×10^5 volts. If $E = 1.90 \times 10^5$ volt/meter, compute the value of B in part (*a*). (*d*) If in part (*c*) the electric field **E** is turned off, determine the radius r of the circle in which the proton now moves.

5. A certain galvanometer has a resistance of 75.3 ohms; its needle experiences a full-scale deflection when a current 1.62×10^{-3} amp passes through its coil. (*a*) Determine the value of the auxiliary resistance required to convert the galvanometer into a voltmeter that reads 1.000 volt at full-scale deflection. How is it to be connected? (*b*) Determine the value of the auxiliary resistance required to convert the galvanometer into an ammeter that reads 0.0500 amp at full-scale deflection. How is it to be connected?

6. In a Hall effect experiment a current of 3.0 amp lengthwise in a conductor 1.0 cm wide, 4.0 cm long, and 10^{-3} cm thick produced a transverse Hall voltage (across the width) of 1.0×10^{-5} volt when a magnetic field of 1.5 weber/meter^2 passed perpendicularly through the thin conductor. From these data, find (*a*) the drift velocity of the charge carriers and (*b*) the number of carriers per cubic centimeter. (*c*) Show on a diagram the polarity of the Hall voltage with a given current and magnetic field direction, assuming the charge carriers are (negative) electrons.

7. (*a*) What is the cyclotron frequency of an electron with an energy of 100 ev in the earth's magnetic field of 1.0×10^{-4} webers/meters^2? (*b*) What is the radius of curvature of the path of this electron if its velocity is perpendicular to the magnetic field?

Chapter 34

1. Eight wires cut the page perpendicularly at the points shown in Fig. 34–32. A wire labeled with the integer k ($k = 1, 2, \ldots 8$) bears the current ki_0. For those with odd k, the current flows up out of the page; for those with even k it flows down into the page. Evaluate $\oint \mathbf{B} \cdot d\mathbf{l}$ along the closed path shown in the direction indicated by the single arrowhead.

2. Two long straight wires pass near one another at right angles. If the wires are free to move, describe what happens when currents are sent through them.

3. Suppose, in Fig. 34–25, that the currents are all in the same direction. What is the force per meter (magnitude and direction) on any one wire? In the analogous case of parallel motion of charged particles in a plasma this is known as the pinch effect.

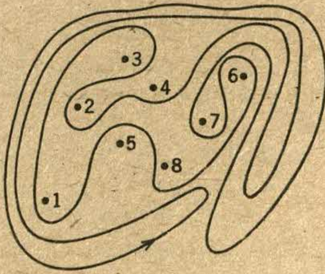


Fig. 34–32

4. A long straight conductor has a circular cross section of radius R and carries a current i. Inside the conductor there is a cylindrical hole of radius a whose axis is parallel to the conductor axis at a distance b from it. Use superposition ideas, and obtain an expression for the magnetic induction B inside the hole.

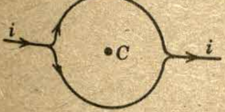
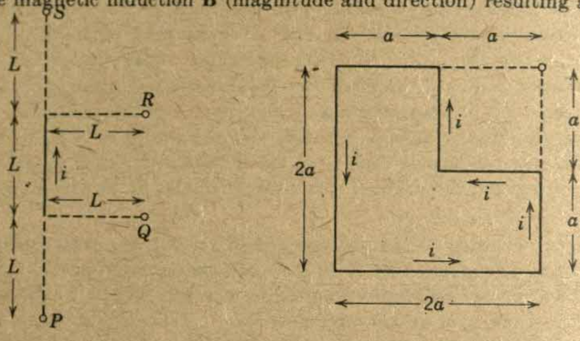


Fig. 34–33

5. A straight conductor is split into identical semicircular turns as shown in Fig. 34–33. What is the magnetic field at the center C of the circular loop so-formed?

6. (*a*) A current i flows in a straight wire of length L in the direction shown in Fig. 34–34*a*. Starting from the Biot-Savart law, determine the resulting magnetic induc-

tion ($\mathbf{B}_P$, $\mathbf{B}_Q$, $\mathbf{B}_R$, $\mathbf{B}_S$, respectively-direction and magnitude in each case) at each of the four points P, Q, R, S (all coplanar with the wire). (*b*) Using the results of part (*a*), compute the magnetic induction **B** (magnitude and direction) resulting at the point T from the current flowing as indicated in the six-sided rectilinear closed loop shown in Fig. 34–34*b*. (Everything drawn is meant to lie in the same plane and all angles are 90°.)



(*a*) (*b*)

Fig. 34–34

7. Use the Biot-Savort law to calculate the magnetic induction **B** at C, the common center of the semicircular arcs AD and HJ, of radii R_2 and R_1 respectively, forming part of the circuit $ADJHA$ carrying current i, as shown in Fig. 34–35.

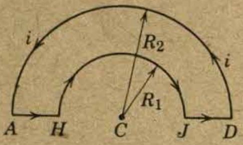


Fig. 34–35

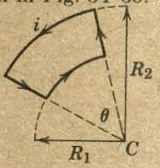


Fig. 34–36

8. Compute the magnetic induction **B** at C, the common center of the circular arcs of the circuit carrying current i, the arcs cutting a sector of angle θ, as shown in Fig. 34–36.

9. (*a*) A long wire is bent into the shape shown in Fig. 34–37, without cross-contact at P. Determine the magnitude and direction of **B** at the center C of the circular portion when the current i flows as indicated. (*b*) The circular part of the wire is rotated without distortion about its (dashed) diameter perpendicular to the straight portion of the wire. The magnetic moment associated with the circular loop is now in the direction of the current in the straight part of the wire. Determine **B** at C in this case.

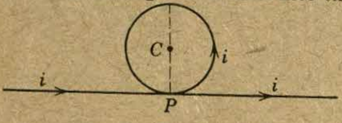


Fig. 34–37

Chapter 35

1. Show that emf has the same dimensions as time rate of change of magnetic flux.

2. In Fig. 35–39, $l = 2.0$ meters and $v = 50$ cm/sec. **B** is the earth's magnetic field, directed perpendicularly out of the page and having a magnitude 6.0×10^{-5} webers/m^2 at that place. The resistance of the circuit $ADCB$, assumed constant (explain how this may be achieved approximately), is $R = 1.2 \times 10^{-5}$ ohms. (*a*) What is the emf

induced in the circuit? (*b*) What is the electric field strength in the wire AB? (*c*) What force does each electron in the wire experience due to the motion of the wire in the magnetic field? (*d*) What is the magnitude and direction of the current in the wire? (*e*) What force must an external agency exert in order to keep the wire moving with this constant velocity? (*f*) Compute the rate at which the external agency is doing work. (*g*) Compute the rate at which electrical energy is being converted into heat energy.

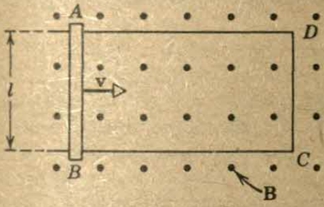


Fig. 35–39

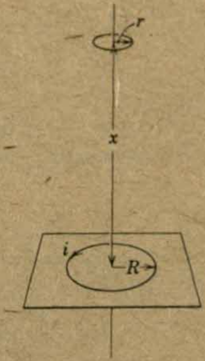
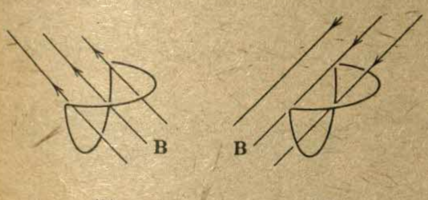


Fig. 35–40

3. Figure 35–40 shows two loops of wire having the same axis. The smaller loop is above the larger one, a distance x which is large compared to the radius R of the larger loop. Hence, with current i flowing as indicated in the larger loop, the consequent magnetic induction is nearly constant throughout the plane area πr^2 bounded by the smaller loop. Suppose now that x is not constant but is changing at the constant rate $dx/dt = v$. (*a*) Determine the magnetic flux across the area bounded by the smaller loop as a function of x. (*b*) Compute the emf generated in the smaller loop at the instant when $x = NR$. (*c*) Determine the direction of the induced current flowing in the smaller loop if $v > 0$.

4. A closed loop of wire consists of a pair of equal semicircles, radius 3.70 cm, lying in mutually perpendicular planes. The loop was formed by folding a circular loop along a diameter until the two halves became perpendicular. A uniform magnetic field **B** of magnitude 760 gauss is directed perpendicular to the fold diameter and makes equal angles (45°) with the planes of the semicircles as shown in Fig. 35–41*a*. (*a*) The magnetic field is reduced at a uniform rate to zero during a time interval 4.50×10^{-3} sec. Determine the magnitude of the induced emf and the sense of the induced current in the loop during this interval. (*b*) How would the answers change if **B** is directed as shown in Fig. 35–41*b*, perpendicular to the direction first given for it but still perpendicular to the "fold-diameter?"



(*a*) (*b*)

Fig. 35–41

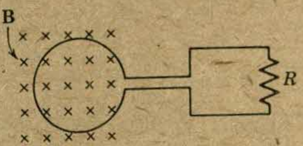


Fig. 35–42

5. A loop of wire of area A is connected to a resistance R. The loop is exposed to a time varying **B** field (see Fig. 35–42). (*a*) Derive an expression for the *net* charge

transferred through the resistor between $t = t_1$ and $t = t_2$. Show that your answer is proportional to the difference $\Phi_B(t_2) - \Phi_B(t_1)$, and is otherwise independent of the manner in which **B** is changing. (*b*) Suppose the change in flux, $\Phi_B(t_2) - \Phi_B(t_1)$ is zero. Does it then follow that no joule-heating occurred during this time interval?

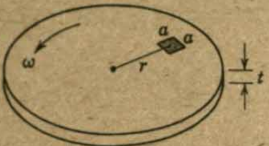


Fig. 35–43

6. An electromagnetic "eddy current" brake consists of a disk of conductivity σ and thickness t rotating about an axis through its center with a magnetic field **B** applied perpendicular to the plane of the disk over a small area a^2 (see Fig. 35–43). If the area a^2 is at a distance r from the axis, find an approximate expression for the torque tending to slow down the disk at the instant its angular velocity equals ω.

Chapter 36

1. Two equal solenoids (inductance L) are wired in series. (*a*) If they are far apart, show that the equivalent inductance is $2L$. (*b*) If they are "close wound," show that the equivalent inductance is either zero or $4L$ depending on the direction of the windings. (*Hint:* They form a single solenoid.)

2. In the circuit shown in Fig. 36–8, $\mathcal{E} = 10$ volts, $R_1 = 5.0$ ohms, $R_2 = 10$ ohms, and $L = 5.0$ henries. For the two separate conditions (I) switch S just closed and (II) switch S closed for a very long time, calculate (*a*) the current i_1 through R_1, (*b*) the current L_2 through R_2, (*c*) the current i through the switch, (*d*) the voltage across R_2, (*e*) the voltage across L, and (*f*) di_2/dt.

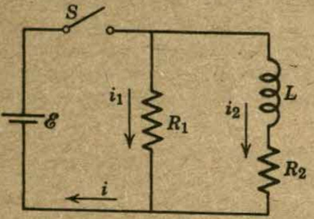


Fig. 36–8

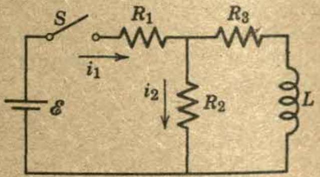


Fig. 36–9

3. In Fig. 36–9, $\mathcal{E} = 100$ volts, $R_1 = 10\,\Omega$, $R_2 = 20\,\Omega$, $R_3 = 30\,\Omega$ and $L = 2$ h. Find the value of i_1 and i_2 (*a*) immediately after S is closed; (*b*) a long time later; (*c*) immediately after S is opened again; (*d*) a long time later.

4. A coaxial cable (inner radius a, outer radius b) is used as a transmission line between a battery $\mathcal{E}$ and a resistor R. The cable acts as both a capacitor and an inductor. If the stored electric and magnetic energies are equal, show that R must be

$$R = \frac{377}{2\pi} \ln\left(\frac{b}{a}\right) \text{ ohms.}$$

(The quantity 377 ohms is sometimes called *the impedance of free space.*)

5. Each item (*a*) coulomb-ohm-meter/weber, (*b*) volt-second, (*c*) coulomb-ampere/farad, (*d*) kilogram-volt-meter^2/(henry-ampere)^2, (*e*) (henry/farad)$^{1/2}$ is equal to one of the items in the following list: meter, second, kilogram, dimensionless number, newton, joule, volt, ohm, watt, coulomb, ampere, weber, henry, farad. Give the equalities.

Chapter 37

1. A total charge q is distributed uniformly on a dielectric ring of radius r. If the ring is rotated about an axis perpendicular to its plane and through its center at an angular speed ω, find the magnitude and direction of its resulting magnetic moment.

2. (*a*) What is the magnetic moment due to the orbital motion of an electron in an atom when the orbital angular momentum is one quantum unit ($= \hbar = 1.05 \times 10^{-34}$ joule-sec). (*b*) The intrinsic spin magnetic moment of an electron is 0.928×10^{-23} amp-meter^2. What is the difference in the magnetic potential energy U between the states in which the magnetic moment is aligned with and aligned in the opposite direction to an external magnetic field of 1.2 weber/meter^2? (*c*) What absolute temperature would be required so that the energy difference in (b) would equal the mean thermal energy $kT/2$?

3. An electron travels in a circular orbit about a fixed positive point charge in the presence of a uniform magnetic field **B** directed normal to the plane of its motion. The electric force has precisely N times the magnitude of the magnetic force on the electron. (*a*) Determine the two possible angular speeds of the electron's motion. (*b*) Evaluate these speeds numerically if $B = 4.27 \times 10^3$ gauss and $N = 100$.

4. A simple bar magnet is suspended by a string as shown in Fig. 37–23. If a uniform magnetic field **B** directed parallel to the ceiling is then established, show the resulting orientation of string and magnet.

5. An iron magnet containing iron of relative permeability 5000 has a flux path 1.0 meter long in the iron and an air gap 0.01 meter long each with cross-sectional areas of 0.02 meter.^{2} What current is necessary in a 500 turn coil wrapped around the iron to give a flux density in the air gap of 1.8 webers/meter^2?

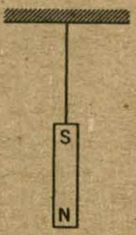


Fig. 37–23

Chapter 38

1. Initially the 900 μf capacitor is charged to 100 volts, and the 100 μf capacitor is uncharged in Fig. 38–13. (*a*) Describe in detail how one may charge the 100 μf capacitor to 300 volts using S_1 and S_2 appropriately. (*b*) Describe in detail the mass + spring mechanical analogy of this problem.

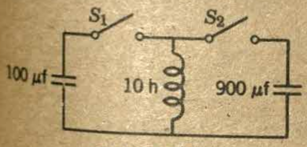


Fig. 38–13

2. *A self-consistency property of two of the Maxwell equations.* Two adjacent parallelepipeds share a common face as shown in Fig. 38–14. (*a*) We may apply $\epsilon_0 \oint \mathbf{E} \cdot d\mathbf{S} = q$ to each of the two closed surfaces separately. Show that, from this alone, it follows that $\epsilon_0 \oint \mathbf{E} \cdot d\mathbf{S} = q$ is *automatically* satisfied for the composite closed surface. (*b*) Repeat using $\oint \mathbf{B} \cdot d\mathbf{S} = 0$.

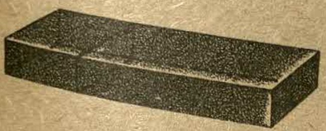


Fig. 38–14

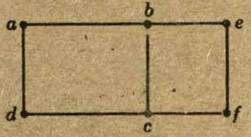


Fig. 38–15

3. *A self-consistency property of two of the Maxwell equations.* Two adjacent closed paths *abcda* and *efcbe* share the common edge *bc* as shown in Fig. 38–15. (*a*) We may apply $\oint \mathbf{E} \cdot d\mathbf{l} = -d\Phi_B/dt$ to each of these closed paths separately. Show that from

this alone, $\oint \mathbf{E} \cdot d\mathbf{l} = -d\Phi_B/dt$ is *automatically* satisfied for the composite closed path *abefcda*. (*b*) Repeat using $\frac{1}{\mu_0} \oint \mathbf{B} \cdot d\mathbf{l} = i + \epsilon_0 \frac{d\Phi_E}{dt}$.

4. Using the definitions of flux, volume charge density ρ, and current density $\mathbf{j}$, write the four Maxwell equations in such a manner that all the fluxes, currents, and charges appear as volume or surface integrals.

Chapter 39

1. A coaxial cable is made of a center wire of radius a surrounded by a thin metal tube of radius b. The substance between the conductors is air. (*a*) Find the capacitance per unit length of this coaxial cable (*Hint:* Imagine equal but opposite charges to be on the wire and the tube). (*b*) Find the inductance per unit length of this coaxial cable (*Hint:* Imagine a current i flowing down the center wire and back along the tube.)

2. Starting with Eqs. 39-10 and 39-13, show that E satisfies the "wave equation"

$$\frac{\partial^2 E}{\partial x^2} = \mu_0 \epsilon_0 \frac{\partial^2 E}{\partial t^2}.$$

Does B also satisfy the wave equation?

3. Show that (*a*) through (*d*) below satisfy Eqs. 39-10 and 39-13. In each of these, A is a constant. (*a*) $E = Ac(x - ct)$, $B = A(x - ct)$. (*b*) $E = Ac(x + ct)^{15}$, $B = -A(x + ct)^{15}$. (*c*) $E = Ace^{(x-ct)}$, $B = Ae^{(x-ct)}$. (*d*) $E = Ac \ln (x + ct)$, $B = -A \ln (x + ct)$. (*e*) Generalize these examples to show that $E = Acf(x - ct)$, $B = Af(x - ct)$ is a solution where f is *any* function of $(x - ct)$. What is the corresponding situation for functions of $(x + ct)$?

4. Consider the possibility of "standing waves":

$$E = E_m(\sin \omega t)(\sin kx)$$
$$B = B_m(\cos \omega t)(\cos kx)$$

(*a*) Show that these satisfy Eqs. 39–10 and 39–13 if E_m is suitably related to B_m and ω suitably related to k. What are these relationships? (*b*) Find the (instantaneous) Poynting vector. (*c*) Show that the time average power flow across any area is zero. (*d*) Describe the flow of energy in this problem.

5. An airplane flying at a distance of 10 km from a radio transmitter receives a signal of power 10.0 microwatts/meter^2. Calculate (*a*) the (average) electric field at the airplane due to this signal; (*b*) the (average) magnetic field at the airplane; (*c*) the total power radiated by the transmitter, assuming the transmitter to radiate isotropically and the earth to be a perfect absorber.

6. A coaxial cable (inner radius a, outer radius b) is used as a transmission line between a battery $\mathcal{E}$ and a resistor R, as shown in Fig. 39–18. (*a*) Calculate **E**, **B** for

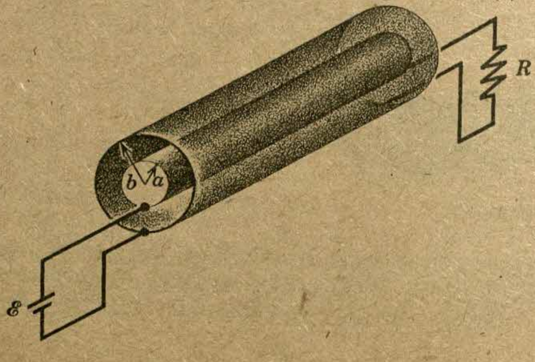


Fig. 39–18

$a < r < b$. (*b*) Calculate the Poynting vector **S** for $a < r < b$. (*c*) By suitably integrating the Poynting vector, show that the total power flowing across the annular cross section $a < r < b$ is $\mathcal{E}^2/R$. Is this reasonable? (*d*) Show that the direction of **S** is always from the battery to the resistor, no matter which way the battery is connected.

7. A long hollow cylinder (radius R, length l) carries a uniform charge per unit area of σ on its surface. An externally applied torque causes the cylinder to rotate at constant acceleration $\omega(t) = \alpha t$. (*a*) Find **B** within the cylinder (treat it as a solenoid). (*b*) Find **E** at the inner surface of the cylinder. (*c*) Find **S** at the inner surface of the cylinder. (*d*) Show that the flux of **S** entering the interior volume of the cylinder is equal to $\frac{d}{dt}\left(\frac{\pi R^2 l}{2\mu_0} B^2\right)$ (Compare with Problem 18.)

Chapter 40

1. A plane electromagnetic wave, with wavelength 3.0 meters, travels in free space in the $+x$-direction with its electric vector **E**, of amplitude 300 volts/meter, directed along the y-axis. (*a*) What is the frequency ν of the wave? (*b*) What is the direction and amplitude of the **B** field associated with the wave? (*c*) If $E = E_m \sin(kx - \omega t)$, what are the values of k and ω for this wave? (*d*) What is the time averaged rate of energy flow per unit area associated with this wave? (*e*) If the wave fell upon a perfectly absorbing sheet of area A, what momentum would be delivered to the sheet per second and what is the radiation pressure exerted on the sheet?

2. Show that $\epsilon_0 \mathbf{E} \times \mathbf{B}$ has the dimensions of momentum/volume. (The vector $\epsilon_0 \mathbf{E} \times \mathbf{B}$ may be used to compute the momentum stored in the fields in the same manner that the scalar $\frac{1}{2}\epsilon_0 \mathbf{E}^2 + \frac{1}{2\mu_0}\mathbf{B}^2$ may be used to compute the energy stored in the fields.)

3. A particle in the solar system is under the combined influence of the sun's gravitational attraction and the radiation force due to the sun's rays. Assume that the particle is a sphere of density 1.0 gm/cm^3 and that all of the incident light is absorbed. (*a*) Show that all particles with radius less than some critical radius, R_0, will be blown out of the solar system. (*b*) Calculate R_0. (*c*) Does R_0 depend on the distance from the earth to the sun? (See the appendices for the necessary constants.)

Chapter 41

1. When an electron moves through a medium at a speed exceeding the speed of light in that medium, it radiates electromagnetic energy (the Cerenkov effect, see Section 20–7). What minimum speed must an electron have in a liquid of refractive index 1.54 in order to radiate?

2. Assume that the index of refraction of the earth's atmosphere varies, with altitude only, from the value one at the edge of the atmosphere to some larger value at the surface of the earth. (*a*) Neglecting the earth's curvature, show that the apparent angle of a star from the zenith direction is independent of how the refractive index of the atmosphere varies with altitude and depends only on the value of n at the earth's surface. (*Hint:* Compare a uniform atmosphere with one consisting of layers of increasing refractive index.) (*b*) How does the earth's curvature affect the analysis?

3. A point source is 80 cm below the surface of a body of water. Find the diameter of the largest circle at the surface through which light can emerge from the water.

4. A pole extends 2.0 meters above the bottom of a swimming pool and 0.5 meters above the water. Sunlight is incident at 45°. What is the length of the shadow of the pole on the bottom of the pool?

5. A given monochromatic light ray, initially in air, strikes the 90° prism at P (see Fig. 41–24) and is refracted there and at Q to such an extent that it just grazes the right-hand prism surface after it emerges into air at Q. (*a*) Determine the index of

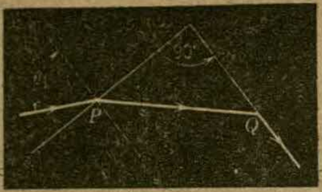


Fig. 41–24

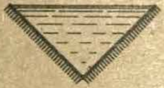


Fig. 41–25

refraction, relative to air, of the prism for this wavelength in terms of the angle of incidence θ_1 which gives rise to this situation. (*b*) Give a numerical upper bound for the index of refraction of the prism. (*c*) Show, by a ray diagram, what happens if the angle of incidence at P is slightly greater than θ_1, is slightly less than θ_1.

6. Two perpendicular mirrors form the sides of a vessel filled with water, as shown in Fig. 41–25. A light ray is incident from above, normal to the water surface. (*a*) Show that the emerging ray is parallel to the incident ray. Assume that there are two reflections at the mirror surfaces. (*b*) Repeat the analysis for the case of oblique incidence, the ray lying in the plane of the figure. (*c*) Using three mirrors, state and prove the three-dimensional analog to this problem.

Chapter 42

1. How many images of himself can an observer see in a room whose ceiling and two adjacent walls are mirrors? Explain.

2. A thin flat plate of glass is a distance b from a convex mirror. A point source of light S is placed a distance a in front of the plate (see Fig. 42–33) so that its image in the partially reflecting plate coincides with its image in the mirror. If $b = 7.5$ cm and the focal length of the mirror is $f = -30$ cm, find a and draw the ray diagram.

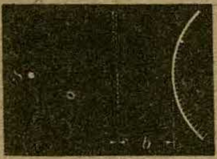


Fig. 42–33

3. (*a*) A simple magnifier (a converging lens) is used directly in front of the eye to examine an object. Draw a ray diagram showing the final image on the retina of the eye. (*b*) The smallest distance at which a normal eye can focus upon an object is 25 cm. Use this result with your ray diagram to prove that the (angular) magnification of a simple magnifier is $25/f + 1$, where f is the focal length of the magnifier lens.

4. Two thin lens, one having $f = +12.0$ cm and the other $f = -10.0$ cm, are separated by 7.0 cm. A small object is placed 43.5 cm from the center of the lens system on the principal axis first on one side and next on the other side. Find the location of the final image in each case.

5. Show that a thin converging lens of focal length f followed by a thin diverging lens of focal length $-f$ will bring parallel light to a focus beyond the second lens provided that the separation of the lenses L satisfies $0 < L < f$. (*b*) Does this property change if the lenses are interchanged? (*c*) What happens when $L = 0$?

6. A concave and a convex lens are cut out of a plane parallel block of glass as shown in Fig. 42–34. Discuss the geometry of a beam of parallel rays incident (*a*) on the concave lens and (*b*) on the convex lens as the distance between lenses is increased from contact to a large separation.

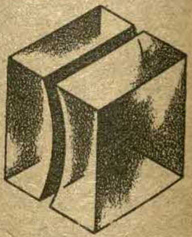


Fig. 42–34

7. An object is placed 1.0 meter in front of a convergent lens, of focal length 0.5 meter, which is 2.0 meters in front of a plane mirror. (*a*) Where is the final image, measured from the lens, that would be seen by an eye looking toward the mirror through the lens? (*b*) Is the final image real or virtual? (*c*) Is the final image erect or inverted? (*d*) What is the lateral magnification?

Chapter 43

1. In a double-slit arrangement the distance between slits is 5.0 mm and the slits are 1.0 meter from the screen. Two interference patterns can be seen on the screen, one due to light of 4800 A and the other 6000 A. What is the separation on the screen between the third-order interference fringes of the two different patterns?

2. In Young's interference experiment in a large ripple tank (see Fig. 43–4) the coherent vibrating sources are placed 12.0 cm apart. The distance between maxima 2.0 meters away is 18.0 cm. If the speed of ripples is 25.0 cm/sec, find the frequency of the vibrators.

3. One slit of a double-slit arrangement is covered by a thin glass plate of refractive index 1.4, and the other by a thin glass plate of refractive index 1.7. The point on the screen where the central maximum fell before the glass plates were inserted is now occupied by what had been the fifth bright fringe before. Assume $\lambda = 4800$ A and that the plates have the same thickness t and find the value of t.

4. In Fig. 43–25, the source emits monochromatic light of wavelength λ. S is a very narrow slit in an otherwise opaque screen I. A plane mirror, whose surface includes the axis of the lens shown, is located a distance h below S. Screen II is at the focal plane of the lens. (*a*) Find the condition for maxima and minima brightness of fringes on screen II in terms of the usual angle θ, the wavelength λ, and the distance h. (*b*) Do fringes appear only in region A (above the axis of the lens), only in region B (below the axis of the lens), or in both regions A and B? Explain. (*Hint:* Consider the image of S formed by the mirror.)

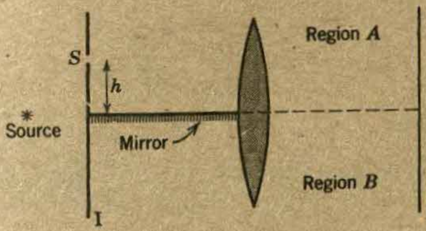


Fig. 43–25

5. Let $f_1(t) = A_1 \sin(\omega t + \phi_1)$ and $f_2(t) = A_2 \sin(\omega t + \phi_2)$. Suppose we want to calculate the time average of their product:

$$\frac{1}{T}\int_{t=0}^{T} f_1(t) f_2(t)\, dt, \qquad T = \frac{2\pi}{\omega}.$$

(Such a problem might arise if f_1 represents a current and f_2 a voltage; $f_1 f_2$ would then be the *instantaneous* power.) Show that this average is equal to one-half the dot product of the corresponding phasors.

6. Consider the problem of determining the sum

$$A_1 \sin(\omega t + \phi_1) + A_2 \sin(\omega t + \phi_2) + \cdots + A_n \sin(\omega t + \phi_n)$$

from the phasor diagram.

(*a*) Show that the sum may always be written in the form

$$B \sin \omega t + C \cos \omega t.$$

(*b*) Show that $B^2 + C^2 \leq (A_1 + A_2 + \cdots + A_n)^2$.
(*c*) When does the equality sign in (*b*) hold?

7. A thin film of acetone (refractive index 1.25) is floated on a thick glass plate (refractive index 1.50). Plane light waves of variable wavelength are incident normal to the film. When one views the reflected wave it is noted that complete destructive interference occurs at 6000 A and constructive interference at 7000 A. Calculate the thickness of the acetone film.

8. We wish to coat a flat piece of glass ($n = 1.50$) with a transparent material ($n = 1.25$) so that light of wavelength 6000 A (in vacuum) incident normally is not reflected. How can this be done?

9. Light of wavelength 6300 A is incident normally on a thin wedge-shaped film of index of refraction 1.5. There are ten bright and nine dark fringes over the length of film. By how much does the film thickness change over this length?

10. An oil drop ($n = 1.20$) floats on a water ($n = 1.33$) surface and is observed from above by reflected light (see Fig. 43–26). (*a*) Will the outer (thinnest) regions of the drop correspond to a bright or a dark region? (*b*) Approximately how thick is the oil film where one observes the third blue region from the outside of the drop? (*c*) Why do the colors gradually disappear as the oil thickness becomes larger?

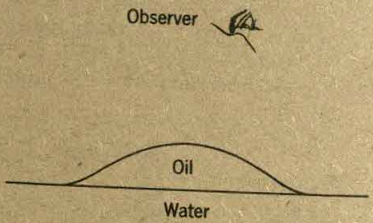


Fig. 43–26

Chapter 44

1. It can be shown that, except for $\theta = 0$, a circular obstacle produces the same diffraction pattern as a circular hole of the same diameter. Furthermore, if there are many such obstacles located randomly, then the interference effects vanish leaving only the diffraction associated with a *single* obstacle. (*a*) Explain why one sees a "ring" around the moon on a foggy day. (*b*) Calculate the size of the water droplet in the air if the bright ring around the moon appears to have a diameter 1.5 times that of the moon.

2. A monochromatic beam of parallel light is incident on a "collimating" hole of diameter $x \gg \lambda$. Point P lies in the geometrical shadow region on a distance screen, as shown in Fig. 44–15*a*. Two obstacles, shown in Fig. 44–15*b*, are placed in turn over the collimating hole. A is an opaque circle with a hole in it and B is the "photographic negative" of A. Using superposition concepts, show that the intensity at P is identical for each of the two diffracting objects A and B (*Babinet's principle*).

3. An astronaut in a satellite claims he can just barely resolve two point sources on the earth, 100 miles below him. What is their separation, assuming ideal conditions? Take $\lambda = 5500$ A, and the pupil diameter to be 5.0 mm.

4. Under ideal conditions, estimate the linear separation of two objects on the planet Mars which can just be resolved by an observer on earth (*a*) using the naked eye,

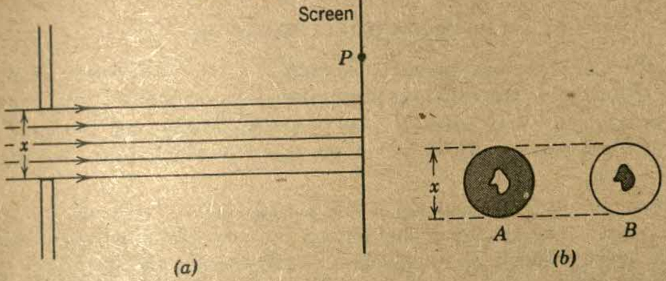


Fig. 44–15

(*b*) using the 200-in. Mt. Palomar telescope. Use the following data: distance to Mars = 50 million miles; diameter of pupil = 5.0 mm; wavelength of light = 5500 A.

5. A double-slit system (slit separation d, slit width a) is driven by two loudspeakers as shown in Fig. 44–16. By use of a variable delay line, the phase of one of the speakers may be varied. Describe in detail what changes occur in the intensity pattern at large distances as this phase difference is varied from zero to 2π. Take both the interference and diffraction effects into account.

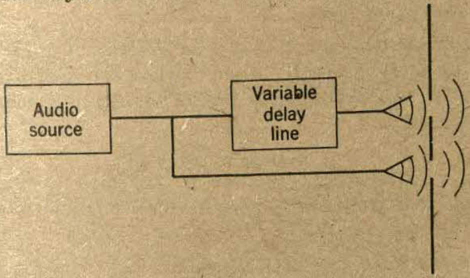


Fig. 44–16

Chapter 45

1. A diffraction grating has a large number N of slits, each of width d. Let $I_{\max}$ denote the intensity at some principal maximum, and let I_k denote the intensity of the kth adjacent secondary maxima. (*a*) If $k \ll N$, show from the phasor diagram that, approximately, $I_k/I_{\max} = 1/(k + \frac{1}{2})^2\pi^2$. (Compare this with the single-slit formula.) (*b*) For those secondary maxima which lie roughly midway between two adjacent principal maxima, show that roughly $I_k/I_{\max} = 1/N^2$. (*c*) Consider the central principal maximum and those adjacent secondary maxima for which $k \ll N$. Show that this part of the diffraction pattern quantitatively resembles that of one single slit of width Nd.

2. With light from a gaseous discharge tube incident normally on a grating with a distance 1.732×10^{-4} cm between adjacent slit centers, a green line appears with sharp maxima at measured transmission angles $\theta = \pm 17.6°, 37.3°, -37.1°, 65.2°$, and $-65.0°$. (*a*) Compute the wavelength of the green line that best fits the data.

3. A narrow beam of monochromatic light strikes a grating at normal incidence and produces sharp maxima at the following angles from the normal: 6°40′, 13°30′, 20°20′, 35°40′. No other maxima appear at any angle between 0° and 35°40′. The separation between adjacent slit centers in the grating is 5.04×10^{-4} cm. (*a*) Compute the wavelength of the light used. (*b*) Make the most complete *quantitative* statement that can be inferred from the above data concerning the width of each slit.

4. Two spectral lines have wavelengths λ and $\lambda + \Delta\lambda$, respectively, where $\Delta\lambda \ll \lambda$. Show that their angular separation $\Delta\theta$ in a grating spectrometer is given approximately by $\Delta\theta = \Delta\lambda/\sqrt{(d/m)^2 - \lambda^2}$, where d is the separation of adjacent slit centers and m is the order at which lines are observed. Notice that the angular separation is greater in the higher orders.

5. A diffraction grating is made up of slits of width 3000 A with a 9000 A separation between centers. The grating is illuminated by monochromatic plane waves, $\lambda =$ 6000 A, the angle of incidence being zero. (*a*) How many diffracted lines are there? (*b*) What is the angular width of the spectral lines observed, if the grating has 1000 slits? (*c*) How is the angular width of the spectral lines related to the resolving power of the grating?

6. A diffraction grating has a resolving power $R = \lambda/\Delta\lambda = Nm$. (*a*) Show that the corresponding frequency range, $\Delta\nu$, that can just be resolved is given by $\Delta\nu = c/Nm\lambda$. (*b*) From Fig. 45–1, show that the "times of flight" of the two extreme rays differ by an amount $\Delta t = Nd \sin\theta/c$. (*c*) Show that $(\Delta\nu)(\Delta t) = 1$, this relation being independent of the various grating parameters.

7. Consider an infinite two-dimensional square lattice as in Fig. 45–14*b*. One interplanar spacing is obviously a_0 itself. (*a*) Calculate the next five smaller interplanar spacings by sketching figures similar to Figure 45–16*a*. (*b*) Show that your answers obey the general formula

$$d = a_0/\sqrt{h^2 + k^2}$$

where h, k are both relatively prime integers (no common integer factor other than unity).

Chapter 46

1. A beam of light is a mixture of plane polarized light and randomly polarized light. When it is sent through a Polaroid sheet, it is found that the transmitted intensity can be varied by a factor of five depending on the orientation of the Polaroid. Find the relative intensities of these two components of the incident beam.

2. It is desired to rotate the plane of polarization of a beam of plane polarized light by 90°. (*a*) How might this be done using only Polaroid sheets? (*b*) How many sheets are required in order that the total intensity loss is less than 5%? Assume that each Polaroid sheet is ideal.

3. A sheet of Polaroid and a quarter-wave plate are glued together in such a way that, if the combination is placed with face *A* against a shiny coin, the face of the coin can be seen when illuminated with light of appropriate wavelength. When the combination is placed with face *A* away from the coin, the coin cannot be seen. Which component is on face *A* and what is the relative orientation of the components?

4. A beam of right circularly polarized light is reflected from a mirror. (*a*) Is the reflected beam right or left circularly polarized? (*b*) Has the direction of the associated linear momentum of the light changed? (*c*) Has the direction of the associated angular momentum of light changed? (*d*) Describe the reaction "felt" by the mirror.

Definition of Standards and Fundamental and Derived Physical Constants*

APPENDIX A

The definition of primary standards is by agreement within the International Conference of Weights and Measures whose most recent general meeting was October, 1964 in Paris. Measured and derived values of the fundamental physical constants summarize hundreds of physical measurements made over the years by scientists in all parts of the world. They have been subjected to exhaustive statistical analysis and, with their accompanying error limits (which are given as three standard deviations), represent the best values to date (1963). For most problem work three significant figures suffice, and the "computational" (rounded) values may be used. The data presented are based largely on values given in the *National Bureau of Standards Technical News Bulletin*, Vol. 47, No. 10 (October 1963).

---

* See "A Pilgrim's Progress in Search of the Fundamental Constants," by J. W. M. Du Mond, *Physics Today*, October 1965.

Definition of Standards and Equivalents

Standard	Abbreviation	Equivalent
Meter	m	1,650,763.73 wavelengths in vacuo of the unperturbed transition $2p_{10}$—$5d_5$ in Kr^{86}
Kilogram	kg	mass of the international kilogram at Sèvres, France
Second	sec	9,192,631,770 vibrations of the unperturbed hyperfine transition 4,0—3,0 of the fundamental state $^2S_{\frac{1}{2}}$ in Cs^{133}[a]
Degree Kelvin	°K	defined in the thermodynamic scale by assigning 273.16 °K to the triple point of water
Unified atomic mass unit	amu	$\frac{1}{12}$ the mass of an atom of the C^{12} nuclide
Mole	mol	amount of substance containing the same number of atoms as 12 gm (exactly) of pure C^{12}
Standard acceleration of free fall	g_n	9.80665 meter/sec^2
Normal atmospheric pressure	atm	101,325 nt/meter^2
Thermochemical calorie	cal	4.1840 joules
Liter	*l*	0.001,000,028 meter^3
Inch	in.	0.0254 meter
Pound (avdp.)	lb	0.453,592,37 kg

---

[a] There is no measurable difference between this and the previous standard of time, 1/31,556,925.9747 of the tropical year at 12^h ET, 0 January 1900. For this reason and because even more accurate maser standards may soon be available, the Cs standard was adopted provisionally rather than "permanently."

FUNDAMENTAL AND DERIVED CONSTANTS

Name	Symbol	Computational Value	Best Experimental Value[b]
Speed of light	c	3.00×10^{8} meters/sec	2.997925 ± 0.000003
Permeability constant	μ_0	1.26×10^{-6} henry/meter	$4\pi \times 10^{-7}$ exactly
Permittivity constant	ϵ_0	8.85×10^{-12} farad/meter	8.85418 ± 0.00002
Elementary charge	e	1.60×10^{-19} coul	1.60210 ± 0.00007
Avogadro constant	N_0	6.02×10^{23}/mole	6.02252 ± 0.00028
Electron rest mass	m_e	9.11×10^{-31} kg	9.1091 ± 0.0004
Proton rest mass	m_p	1.67×10^{-27} kg	1.67252 ± 0.00008
Neutron rest mass	m_n	1.67×10^{-27} kg	1.67482 ± 0.00008
Faraday constant	F	9.65×10^{4} coul/mole	9.64870 ± 0.00016
Planck constant	h	6.63×10^{-34} joule sec	6.6256 ± 0.0005
Fine structure constant	α	7.30×10^{-3}	7.29720 ± 0.00010
Electron charge/mass ratio	e/m_e	1.76×10^{11} coul/kg	1.758796 ± 0.000019
Quantum/charge ratio	h/e	4.14×10^{-15} joule sec/coul	4.13556 ± 0.00012
Electron Compton wavelength	λ_C	2.43×10^{-12} meter	2.42621 ± 0.00006
Proton Compton wavelength	λ_{Cp}	1.32×10^{-15} meter	1.32140 ± 0.00004
Rydberg constant	R_∞	1.10×10^{7}/meter	1.0973731 ± 0.0000003
Bohr radius	a_0	5.29×10^{-11} meter	5.29167 ± 0.00007
Bohr magneton	μ_B	9.27×10^{-24} joule/tesla[a]	9.2732 ± 0.0006
Nuclear magneton	μ_N	5.05×10^{-27} joule/tesla[a]	5.0505 ± 0.0004
Proton magnetic moment	μ_p	1.41×10^{-26} joule/tesla[a]	1.41049 ± 0.00013
Universal gas constant	R	8.31 joule/°K mole	8.3143 ± 0.0012
Standard volume of ideal gas	—	2.24×10^{-2} meter^3/mole	2.24136 ± 0.00030
Boltzmann constant	k	1.38×10^{-23} joule/°K	1.38054 ± 0.00018
First radiation constant $\pi 2 hc^2$	c_1	3.74×10^{-16} watt/meter^2	3.7405 ± 0.0003
Second radiation constant hc/k	c_2	1.44×10^{-2} meter °K	1.43879 ± 0.00019
Wien displacement constant	b	2.90×10^{-3} meter °K	2.8978 ± 0.0004
Stefan-Boltzmann constant	σ	5.67×10^{-8} watt/meter^2 °K^4	5.6697 ± 0.0029
Gravitational constant	G	6.67×10^{-11} nt meter^2/kg^2	6.670 ± 0.015

[a] Tesla ≡ weber/meter^2.

[b] Same units and power of ten as the computational value.

Miscellaneous Terrestrial Data

APPENDIX B

Standard atmosphere	1.013×10^5 nt/meter^2 14.70 lb/in^2 760.0 mm-Hg
Density of dry air at STP[a]	1.293 kg/meter^3 2.458×10^{-3} slug/ft^3
Speed of sound in dry air at STP	331.4 meters/sec 1089 ft/sec 742.5 miles/hr
Acceleration of gravity, g (standard value)[b]	9.80665 meters/sec^2 32.1740 ft/sec^2
Solar constant[c]	1340 watts/m^2 1.92 cal/cm^2-min
Mean total solar radiation	3.92×10^{26} watts
Equatorial radius of earth	6.378×10^6 meters 3963 miles
Polar radius of earth	6.357×10^6 meters 3950 miles
Volume of earth	1.087×10^{21} meter^3 3.838×10^{22} ft^3
Radius of sphere having same volume	6.371×10^6 meters 3959 miles 2.090×10^7 ft

Mean density of earth	5522 kg/meter^3
Mass of earth	5.983×10^{24} kg
Mean orbital speed of earth	29,770 meters/sec 18.50 miles/sec
Mean angular speed of rotation of earth	7.29×10^{-5} radians/sec
Earth's magnetic field, B (at Washington, D. C.)	5.7×10^{-5} tesla[d]
Earth's magnetic dipole moment	6.4×10^{21} amp-m^2

---

[a] STP = standard temperature and pressure = 0° C and 1 atm.

[b] This value, used for barometer corrections, legal weights, etc., was adopted by the International Committee on Weights and Measures in 1901. It approximates 45° latitude at sea level.

[c] The solar constant is the solar energy falling per unit time at normal incidence on unit area of the earth's surface.

[d] Tesla ≡ weber/meter^2.

The Solar System[a]

APPENDIX C

Planet	Mercury ☿	Venus ♀	Earth ⊕, ⊖, ♁	Mars ♂	Jupier ♃	Saturn ♄	Uranus ⛢, ♅	Neptune ♆	Pluto ♇
Mean diameter km	5,000	12,400	12,742	6,870	139,760	115,100	51,000	50,000	12,700?
Earth diameters	0.39	0.973	1.000	0.532	10.97	9.03	4.00	3.90	0.46
Volume (earth volumes)	0.06	0.92	1.00	0.15	1,318	736	64	39	0.10
Mass (earth masses)	0.04	0.82	1.00	0.11	318.3	95.3	14.7	17.3	1.0?
Density (earth densities)	0.69	0.89	1.00	0.70	0.24	0.13	0.23	0.29	?
Mean density gm/cm^3	3.8	4.86	5.52	3.96	1.33	0.71	1.26	1.6	?
Surface gravity (earth's)	0.27	0.86	1.00	0.37	2.64	1.17	0.92	1.44	?
Velocity of escape, km/sec	3.6	10.2	11.2	5.0	60	36	21	23	11?
Length of day (earth days)	58.6^d	30^d?	1^d	1^{d}37^{m}23^s	9^{h}55^m	10^{h}38^m	10.7^h	15.8^h	?

Planet	Mercury ☿	Venus ♀	Earth, ⊕, ⊖, ♁	Mars ♂	Jupiter ♃	Saturn ♄	Uranus ⛢, ♅	Neptune ♆	Pluto ♇
Period, siderial days	87.97	224.70	365.26	686.98	4,332.59	10,759.20	30,685.93	60,187.64	90,885
Inclination of equator to orbit	—	0°?	23°27′	25°12′	3°7′	26°45′	98.0°	29°	?
Oblateness	0.00	0.00	1/296	1/192	1/15.4	1/9.5	1/14	1/45	?
Atmosphere, main constituents	none	N_2, CO_2, A	N_2, O_2	N_2, CO_2, H_2O	CH_4, NH_3	CH_4, NH_3	CH_4, NH_3	CH_4, NH_3	none
Maximum surface temperature, °K	700	700	350	320	153	138	110?	90?	80?
Distance from Sun, 10^6 km	58	108	149	228	778	1426	2869	4495	5900

The Sun ⊙ 329,390 earth masses, mean density 1.42, mean diameter 1,390,600 km, surface gravity 28 (earth's).

The Moon ☽ 0.01228 earth masses, mean density 3.36, mean diameter 3,476 km, surface gravity 0.17 (earth's), distance from earth 38×10^4 km.

[a] Adapted from Payne-Gaposchkin and *Handbook of Chemistry and Physics.*

Periodic Table of the Elements

APPENDIX D

Atomic weights are expressed in *atomic mass units* (amu), one atom of the isotope C^{12} being defined to have a mass of (exactly) 12 amu. For unstable elements the mass number of the most stable or best known isotope is given in brackets.

Group →		I	II	III	IV	V	VI	VII	VIII			0
Period	Series											
1	1	1 H 1.00797										2 He 4.0026
2	2	3 Li 6.939	4 Be 9.0122	5 B 10.811	6 C 12.01115	7 N 14.0067	8 O 15.9994	9 F 18.9984				10 Ne 20.183
3	3	11 Na 22.9898	12 Mg 24.312	13 Al 26.9815	14 Si 28.086	15 P 30.9738	16 S 32.064	17 Cl 35.453				18 A 39.948
4	4	19 K 39.102	20 Ca 40.08	21 Sc 44.956	22 Ti 47.90	23 V 50.942	24 Cr 51.996	25 Mn 54.9380	26 Fe 55.847	27 Co 58.9332	28 Ni 58.71	
	5	29 Cu 63.54	30 Zn 65.37	31 Ga 69.72	32 Ge 72.59	33 As 74.9216	34 Se 78.96	35 Br 79.909				36 Kr 83.80
5	6	37 Rb 85.47	38 Sr 87.62	39 Y 88.905	40 Zr 91.22	41 Nb 92.906	42 Mo 95.94	43 Tc [99]	44 Ru 101.07	45 Rh 102.905	46 Pd 106.4	
	7	47 Ag 107.870	48 Cd 112.40	49 In 114.82	50 Sn 118.69	51 Sb 121.75	52 Te 127.60	53 I 126.9044				54 Xe 131.30
6	8	55 Cs 132.905	56 Ba 137.34	57–71 Lanthanide series[a]	72 Hf 178.49	73 Ta 180.948	74 W 183.85	75 Re 186.2	76 Os 190.2	77 Ir 192.2	78 Pt 195.09	
	9	79 Au 196.967	80 Hg 200.59	81 Tl 204.37	82 Pb 207.19	83 Bi 208.980	84 Po [210]	85 At [210]				86 Rn [222]
7	10	87 Fr [223]	88 Ra [226]	89 Actinide series[b]								

[a] **Lanthanide series:**	57 La 138.91	58 Ce 140.12	59 Pr 140.907	60 Nd 144.24	61 Pm [145]	62 Sm 150.35	63 Eu 151.96	64 Gd 157.25	65 Tb 158.924	66 Dy 162.50	67 Ho 164.930	68 Er 167.26	69 Tm 168.934	70 Yb 173.04	71 Lu 174.97
[b] **Actinide series:**	89 Ac [227]	90 Th 232.038	91 Pa [231]	92 U 238.04	93 Np [237]	94 Pu [242]	95 Am [243]	96 Cm [245]	97 Bk [249]	98 Cf [249]	99 Es [254]	100 Fm [252]	101 Md [256]	102 [254]	103 Lw [257]

The Particles of Physics[a]

APPENDIX E

Family name	Particle name	Symbol	Mass	Spin	Strangeness[b]	Charge	Antiparticle[c]	No. of particles	Average[d] lifetime, seconds	Typical decay products
	Photon	γ (gamma ray)	0	1	0	0	Same particle	1	Infinite	—
Electron family	Electron	e^-	1	$\frac{1}{2}$	—	$-e$	e^+	2	Infinite	—
	Electron's neutrino	ν_e	0	$\frac{1}{2}$	—	0	$\overline{\nu}_e$	2	Infinite	—
Muon family	Muon	μ^-	206.77	$\frac{1}{2}$	—	$-e$	μ^+	2	2.212×10^{-6}	$e^- + \overline{\nu}_e + \nu_\mu$
	Muon's neutrino	ν_μ	0(?)	$\frac{1}{2}$	—	0	$\overline{\nu}_\mu$	2	Infinite	—
Mesons	Pion	π^+	273.2	0	0	$+e$	π^-	3	2.55×10^{-8}	$\mu^+ + \nu_\mu$
		π^0	264.2	0	0	0	π^0		1.9×10^{-16}	$\gamma + \gamma$
	Kaon	K^+	966.6	0	+1	$+e$	$\overline{K^+}$	4	1.22×10^{-8}	$\pi^+ + \pi^0$
		K^0	974	0	+1	0	$\overline{K^0}$		1.00×10^{-10} and 6×10^{-8}	$\pi^+ + \pi^-$

Family name	Particle name	Symbol	Mass	Spin	Strangeness[b]	Charge	Antiparticle[c]	No. of particles	Average lifetime seconds	Typical decay products
Baryons	Nucleon	p^+ (proton)	1836.12	$\frac{1}{2}$	0	$+e$	$\overline{p^+}$	4	Infinite	
		n^0 (neutron)	1838.65	$\frac{1}{2}$	0	0	$\overline{n^0}$		1013	$p + e^- + \overline{\nu_e}$
	Lambda particle	Λ^0	2182.8	$\frac{1}{2}$	−1	0	$\overline{\Lambda^0}$	2	$2.51 \times 10\ ^{10}$	$p + \pi^-$
	Sigma particle	Σ^+	2327.7	$\frac{1}{2}$	−1	$+e$	$\overline{\Sigma^+}$		8.1×10^{-11}	$n + \pi^+$
		Σ^-	2340.5	$\frac{1}{2}$	−1	$-e$	$\overline{\Sigma^-}$	6	1.6×10^{-10}	$n + \pi^-$
		Σ^0	2332	$\frac{1}{2}$	−1	0	$\overline{\Sigma^0}$		about 10^{-20}	$\Lambda^0 + \gamma$
	Xi particle	Ξ^-	2580	$\frac{1}{2}$	−2	$-e$	$\overline{\Xi^-}$	4	1.3×10^{-10}	$\Lambda^0 + \pi^-$
		Ξ^0	2570	$\frac{1}{2}$	−2	0	$\overline{\Xi^0}$		about 10^{-10}	$\Lambda^0 + \pi^0$

---

[a] Adapted and modified from *The World of Elementary Particles*, by Kenneth W. Ford.

[b] This is a "quantum number" whose assignment permits an understanding of the inter-relationships of the particles.

[c] Antiparticles have the same mass and spin as the particles but their charges and strangeness numbers are opposite in sign.

[d] The K^0 meson has two different lifetimes; all other particles have only one.

Symbols, Dimensions, and Units for Physical Quantities

APPENDIX F

All units and dimensions are in the mksq (rationalized) system. The primary units can be found by reading kilograms for M, meters for L, seconds for T, and coulombs for Q. The symbols are those used in the text.

In practice, Q is defined in terms of M, L, and T. However, the addition of Q to the traditional M, L, and T enables us to avoid the use of fractional exponents in dimensional considerations. The term 'rationalized' simply means that a factor $1/4\pi$ is separated out of Coulomb's law in order to remove the factor 4π that would otherwise appear in many other formulas in electricity.

Quantity	Symbol	Dimensions	Derived Units
Acceleration	$\mathbf{a}$	LT^{-2}	meters/sec^2
Angular acceleration	$\boldsymbol{\alpha}$	T^{-2}	radians/sec^2
Angular displacement	θ	—	radian
Angular frequency and speed	ω	T^{-1}	radians/sec
Angular momentum	$\mathbf{L}$	ML^2T^{-1}	kg-m^2/sec
Angular velocity	$\boldsymbol{\omega}$	T^{-1}	radians/sec
Area	A, S	L^2	meter^2
Displacement	$\mathbf{r}, \mathbf{d}$	L	meter

Quantity	Symbol	Dimensions	Derived Units
Energy, total	E	ML^2T^{-2}	joule
kinetic	K	ML^2T^{-2}	joule
potential	U	ML^2T^{-2}	joule
Force	$\mathbf{F}$	MLT^{-2}	newton
Frequency	ν	T^{-1}	hertz ≡ cycles/sec
Gravitational field strength	$\mathbf{g}$	LT^{-2}	nt/kg
Gravitational potential	V	L^2T^{-2}	joules/kg
Length	l	L	meter
Mass	m	M	kilogram
Mass density	ρ	ML^{-3}	kg/m³
Momentum	$\mathbf{p}$	MLT^{-1}	kg-m/sec
Period	T	T	second
Power	P	ML^2T^{-3}	watt
Pressure	p	$ML^{-1}T^{-2}$	nt/m²
Rotational inertia	I	ML^2	kg-m²
Time	t	T	second
Torque	$\boldsymbol{\tau}$	ML^2T^{-2}	nt-m
Velocity	$\mathbf{v}$	LT^{-1}	meters/sec
Volume	V	L^3	meter³
Wavelength	λ	L	meter
Work	W	ML^2T^{-2}	joule
Entropy	S	ML^2T^{-2}	joules/K°
Internal energy	U	ML^2T^{-2}	joule
Heat	Q	ML^2T^{-2}	joule
Temperature	T	—	Kelvin degree
Capacitance	C	$M^{-1}L^{-2}T^2Q^2$	farad
Charge	q	Q	coulomb
Conductivity	σ	$M^{-1}L^{-3}TQ^2$	(ohm-meter)$^{-1}$
Current	i	$T^{-1}Q$	ampere
Current density	$\mathbf{j}$	$L^{-2}T^{-1}Q$	amp/meter²
Electric dipole moment	$\mathbf{p}$	LQ	coul-meter
Electric displacement	$\mathbf{D}$	$M^{-2}Q$	coul/meter²
Electric polarization	$\mathbf{P}$	$M^{-2}Q$	coul-meter²
Electric field strength	$\mathbf{E}$	$MLT^{-2}Q^{-1}$	volts/meter
Electric flux	Φ_E	$ML^3T^{-2}Q^{-1}$	volt-meter
Electric potential	V	$ML^2T^{-2}Q^{-1}$	volt
Electromotive force	$\mathcal{E}$	$ML^2T^{-2}Q^{-1}$	volt
Inductance	L	ML^2Q^{-2}	henry
Magnetic dipole moment	$\boldsymbol{\mu}$	$L^2T^{-1}Q$	amp-meter²
Magnetic field strength	$\mathbf{H}$	$L^{-1}T^{-1}Q$	amp-meter
Magnetic flux	Φ_B	$ML^2T^{-1}Q^{-1}$	weber = volt-sec
Magnetic induction	$\mathbf{B}$	$MT^{-1}Q^{-1}$	tesla ≡ webers/meter²
Magnetization	$\mathbf{M}$	$L^{-1}T^{-1}Q$	amp/meter
Permeability	μ	MLQ^{-2}	henrys/meter
Permittivity	ϵ	$M^{-1}L^{-3}T^2Q^2$	farads/meter
Resistance	R	$ML^2T^{-1}Q^{-2}$	ohm
Resistivity	ρ	$ML^3T^{-1}Q^{-2}$	ohm-meter
Voltage	V	$ML^2T^{-2}Q^{-1}$	volt

Conversion Factors[a]

APPENDIX G

Conversion factors for common and not-so-common units may be read off directly from the tables below. For example, 1 degree = 2.778×10^{-3} revolutions, so $16.7° = 16.7 \times 2.778 \times 10^{-3}$ rev. The mks quantities are capitalized in each table.

PLANE ANGLE

	°	′	″	RADIAN	rev
1 degree =	1	60	3600	1.745×10^{-2}	2.778×10^{-3}
1 minute =	1.667×10^{-2}	1	60	2.909×10^{-4}	4.630×10^{-5}
1 second =	2.778×10^{-4}	1.667×10^{-2}	1	4.848×10^{-6}	7.716×10^{-7}
1 RADIAN =	57.30	3438	2.063×10^{5}	1	0.1592
1 revolution =	360	2.16×10^{4}	1.296×10^{6}	6.283	1

1 rev = 2π radians = 360° $1° = 60' = 3600''$

SOLID ANGLE

1 sphere = 4π steradians = 12.57 steradians

---

[a] Adapted in part from G. Shortley and D. Williams, *Elements of Physics*, Prentice-Hall, Englewood Cliffs, New Jersey, Second Edition, 1955.

LENGTH

	cm	METER	km	in.	ft	mile
1 centimeter =	1	10^{-2}	10^{-5}	0.3937	3.281×10^{-2}	6.214×10^{-6}
1 METER =	100	1	10^{-3}	39.37	3.281	6.214×10^{-4}
1 kilometer =	10^5	1000	1	3.937×10^4	3281	0.6214
1 inch =	2.540	2.540×10^{-2}	2.540×10^{-5}	1	8.333×10^{-2}	1.578×10^{-5}
1 foot =	30.48	0.3048	3.048×10^{-4}	12	1	1.894×10^{-4}
1 statute mile =	1.609×10^5	1609	1.609	6.336×10^4	5280	1

1 angstrom (A) = 10^{-10} meter
1 X-unit = 10^{-13} meter
1 micron = 10^{-6} meter
1 millimicron (mμ) = 10^{-9} meter
1 nautical mile = 1852 meters = 1.1508 statute miles = 6076.10 ft

1 light-year = 9.4600×10^{12} km
1 parsec = 3.084×10^{13} km
1 fathom = 6 ft

1 yard = 3 ft
1 rod = 16.5 ft
1 mil = 10^{-3} in.

AREA

	METER²	cm²	ft²	in.²	circ mil
1 SQUARE METER =	1	10^4	10.76	1550	1.974×10^9
1 square centimeter =	10^{-4}	1	1.076×10^{-3}	0.1550	1.974×10
1 square foot =	9.290×10^{-2}	929.0	1	144	1.833×10^8
1 square inch =	6.452×10^{-4}	6.452	6.944×10^{-3}	1	1.273×10^6
1 circular mill =	5.067×10^{-10}	5.067×10^{-6}	5.454×10^{-9}	7.854×10^{-7}	1

1 square mile = 27,878,400 ft² = 640 acres
1 barn = 10^{-28} meter²

1 acre = 43,560 ft²

VOLUME

	METER³	cm³	1	ft³	in.³
1 CUBIC METER =	1	10^6	1000	35.31	6.102×10^4
1 cubic centimeter =	10^{-6}	1	1.000×10^{-3}	3.531×10^{-5}	6.102×10^{-2}
1 liter =	1.000×10^{-3}	1000	1	3.531×10^{-2}	61.02
1 cubic foot =	2.832×10^{-2}	2.832×10^4	28.32	1	1728
1 cubic inch =	1.639×10^{-5}	16.39	1.639×10^{-2}	5.787×10^{-4}	1

1 U. S. fluid gallon = 4 U. S. fluid quarts = 8 U. S. pints = 128 U. S. fluid ounces = 231 in.³
1 British imperial gallon = the volume of 10 lb of water at 62° F = 277.42 in.³
1 liter = the volume of 1 kg of water at its maximum density = 1000.028 cm³

MASS

Note: Those quantities to the right of and below the heavy lines are not mass units at all but are often used as such. When we write, for example,

1 kg "=" 2.205 lb

this means that a kilogram is a *mass* that *weighs* 2.205 pounds. Clearly this "equivalence" is approximate (depending on the value of g) and is meaningful only for terrestrial measurements. Thus, care must be employed when using the factors in the shaded portion of the table.

	gm	KG	slug	amu	oz	lb	ton
1 gram =	1	0.001	6.852×10^{-5}	6.024×10^{23}	3.527×10^{-2}	2.205×10^{-3}	1.102×10^{-6}
1 KILOGRAM =	1000	1	6.852×10^{-2}	6.024×10^{26}	35.27	2.205	1.102×10^{-3}
1 slug =	1.459×10^{4}	14.59	1	8.789×10^{27}	514.8	32.17	1.609×10^{-2}
1 amu =	1.660×10^{-24}	1.660×10^{-27}	1.137×10^{-28}	1	5.855×10^{-26}	3.660×10^{-27}	1.829×10^{-30}
1 ounce (avoirdupois) =	28.35	2.835×10^{-2}	1.943×10^{-3}	1.708×10^{25}	1	6.250×10^{-2}	3.125×10^{-5}
1 pound (avoirdupois) =	453.6	0.4536	3.108×10^{-2}	2.732×10^{26}	16	1	0.0005
1 ton =	$9.072 \times 10^{+5}$	907.2	62.16	5.465×10^{29}	3.2×10^{4}	2000	1

DENSITY

Note: Those quantities to the right or below the heavy line are weight densities and, as such, are dimensionally different from mass densities. Care must be used. (See note for mass table.)

	slug/ft^3	KG/METER^3	gm/cm^3	lb/ft^3	lb/in.^{3}
1 slug per ft^3 =	1	515.4	0.5154	32.17	1.862×10^{-2}
1 KILOGRAM per METER^3 =	1.940×10^{-3}	1	0.001	6.243×10^{-2}	3.613×10^{-5}
1 gram per cm^3 =	1.940	1000	1	62.43	3.613×10^{-2}
1 pound per ft^3 =	3.108×10^{-2}	16.02	1.602×10^{-2}	1	5.787×10^{-4}
1 pound per in.^{3} =	53.71	2.768×10^{4}	27.68	1728	1

TIME

	yr	day	hr	min	SEC
1 year =	1	365.2	8.766×10^{3}	5.259×10^{5}	3.156×10^{7}
1 day =	2.738×10^{-3}	1	24	1440	8.640×10^{4}
1 hour =	1.141×10^{-4}	4.167×10^{-2}	1	60	3600
1 minute =	1.901×10^{-6}	6.944×10^{-4}	1.667×10^{-2}	1	60
1 SECOND =	3.169×10^{-8}	1.157×10^{-5}	2.778×10^{-4}	1.667×10^{-2}	1

SPEED

	ft/sec	km/hr	METER/SEC	miles/hr	cm/sec	knot
1 foot per second =	1	1.097	0.3048	0.6818	30.48	0.5925
1 kilometer per hour =	0.9113	1	0.2778	0.6214	27.78	0.5400
1 METER per SECOND =	3.281	3.6	1	2.237	100	1.944
1 mile per hour =	1.467	1.609	0.4470	1	44.70	0.8686
1 centimeter per second =	3.281×10^{-2}	3.6×10^{-2}	0.01	2.237×10^{-2}	1	1.944×10^{-2}
1 knot =	1.688	1.852	0.5144	1.151	51.44	1

1 knot = 1 nautical mile/hr 1 mile/min = 88 ft/sec = 60 miles/hr

FORCE

Note: Those quantities to the right of and below the heavy lines are not force units at all but are often used as such, especially in chemistry. For instance, if we write

1 gram-force "=" 980.7 dynes,

we mean that a gram-*mass* experiences a *force* of 980.7 dynes in the earth's gravitational field. Thus, care must be employed when using the factors in the shaded portion of the table.

	dyne	NT	lb	pdl	gf	kgf
1 dyne =	1	10^{-5}	2.248×10^{-6}	7.233×10^{-5}	1.020×10^{-3}	1.020×10^{-6}
1 NEWTON =	10^5	1	0.2248	7.233	102.0	0.1020
1 pound =	4.448×10^5	4.448	1	32.17	453.6	0.4536
1 poundal =	1.383×10^4	0.1383	3.108×10^{-2}	1	14.10	1.410×10^{-2}
1 gram-force =	980.7	9.807×10^{-3}	2.205×10^{-3}	7.093×10^{-2}	1	0.001
1 kilogram-force =	9.807×10^5	9.807	2.205	70.93	1000	1

1 kgf = 9.80665 nt 1 lb = 32.17398 pdl

PRESSURE

	atm	dyne/cm²	inch of water	cm Hg	NT/METER²	lb/in.²	lb/ft²
1 atmosphere =	1	1.013×10^6	406.8	76	1.013×10^5	14.70	2116
1 dyne per cm² =	9.869×10^{-7}	1	4.015×10^{-4}	7.501×10^{-5}	0.1	1.450×10^{-5}	2.089×10^{-3}
1 inch of water at 4° C[a] =	2.458×10^{-3}	2491	1	0.1868	249.1	3.613×10^{-2}	5.202
1 centimeter of mercury at 0° C[a] =	1.316×10^{-2}	1.333×10^4	5.353	1	1333	0.1934	27.85
1 NEWTON per METER² =	9.869×10^{-6}	10	4.015×10^{-3}	7.501×10^{-4}	1	1.450×10^{-4}	2.089×10^{-2}
1 pound per in.² =	6.805×10^{-2}	6.895×10^4	27.68	5.171	6.895×10^3	1	144
1 pound per ft² =	4.725×10^{-4}	478.8	0.1922	3.591×10^{-2}	47.88	6.944×10^{-3}	1

[a] Where the acceleration of gravity has the standard value 9.80665 meters/sec².

1 bar = 10^6 dyne/cm² 1 millibar = 10^3 dyne/cm²

ENERGY, WORK, HEAT

The electron volt (ev) is the kinetic energy an electron gains from being accelerated through the potential difference of one volt in an electric field. The Mev is the kinetic energy it gains from being accelerated through a million-volt potential difference.

The last two items in this table are not properly energy units but are included for convenience. They arise from the relativistic mass-energy equivalence formula $E = mc^2$ and represent the energy released if a kilogram or atomic mass unit (amu) is destroyed completely.

Again, care should be used when employing this table.

	Btu	erg	ft-lb	hp-hr	JOULES	cal	kw-hr	ev	Mev	kg	amu
1 British thermal unit =	1	1.055×10^{10}	777.9	3.929×10^{-4}	1055	252.0	2.930×10^{-4}	6.585×10^{21}	6.585×10^{15}	1.174×10^{-14}	7.074×10^{12}
1 erg =	9.481×10^{-11}	1	7.376×10^{-8}	3.725×10^{-14}	10^{-7}	2.389×10^{-8}	2.778×10^{-14}	6.242×10^{11}	6.242×10^{5}	1.113×10^{-24}	670.5
1 foot-pound =	1.285×10^{-3}	1.356×10^{7}	1	5.051×10^{-7}	1.356	0.3239	3.766×10^{-7}	8.464×10^{18}	8.464×10^{12}	1.509×10^{-17}	9.092×10^{9}
1 horsepower-hour =	2545	2.685×10^{13}	1.980×10^{6}	1	2.685×10^{6}	6.414×10^{5}	0.7457	1.676×10^{25}	1.676×10^{19}	2.988×10^{-11}	1.800×10^{16}
1 JOULE =	9.481×10^{-4}	10^{7}	0.7376	$3,725 \times 10^{-7}$	1	0.2389	2.778×10^{-7}	6.242×10^{18}	6.242×10^{12}	1.113×10^{-17}	6.705×10^{9}
1 calorie =	3.968×10^{-3}	4.186×10^{7}	3.087	1.559×10^{-6}	4.186	1	1.163×10^{-6}	2.613×10^{19}	2.613×10^{13}	4.659×10^{-17}	2.807×10^{10}
1 kilowatt-hour =	3413	3.6×10^{13}	2.655×10^{6}	1.341	3.6×10^{6}	8.601×10^{5}	1	2.247×10^{25}	2.270×10^{19}	4.007×10^{-11}	2.414×10^{16}
1 electron volt =	1.519×10^{-22}	1.602×10^{-12}	1.182×10^{-19}	5.967×10^{-26}	1.602×10^{-19}	3.827×10^{-20}	4.450×10^{-26}	1	10^{-6}	1.733×10^{-36}	1.074×10^{-9}
1 million electron volts =	1.519×10^{-16}	1.602×10^{-6}	1.182×10^{-13}	5.967×10^{-20}	1.602×10^{-13}	3.827×10^{-14}	4.450×10^{-20}	10^{6}	1	1.783×10^{-30}	1.074×10^{-8}
1 kilogram =	8.521×10^{13}	8.987×10^{23}	6.629×10^{16}	3.348×10^{10}	8.987×10^{16}	2.147×10^{16}	2.497×10^{10}	5.610×10^{35}	5.810×10^{29}	1	6.025×10^{26}
1 atomid mass unit =	1.415×10^{-13}	1.492×10^{-3}	1.100×10^{-10}	5.558×10^{-17}	1.492×10^{-10}	3.564×10^{-11}	4.145×10^{-17}	$9\ 31 \times 10^{8}$	931.0	1.660×10^{-27}	1

1 m-kgf = 9.807 joules 1 watt-sec = 1 joule = 1 nt-m 1 cm-dyne = 1 erg

POWER

	Btu/hr	ft-lb/min	ft-lb/sec	hp	cal/sec	kw	WATTS
1 British thermal unit per hour =	1	12.97	0.2161	3.929×10^{-4}	7.000×10^{-2}	2.930×10^{-4}	0.2930
1 foot-pound per minute =	7.713×10^{-2}	1	1.667×10^{-2}	3.030×10^{-5}	5.399×10^{-3}	2.260×10^{-5}	2.260×10^{-2}
1 foot-pound per second =	4.628	60	1	1.818×10^{-3}	0.3239	1.356×10^{-3}	1.356
1 horsepower =	2545	3.3×10^{4}	550	1	178.2	0.7457	745.7
1 calorie per second =	14.29	1.852×10^{2}	3.087	5.613×10^{-3}	1	4.186×10^{-3}	4.186
1 kilowatt =	3413	4.425×10^{4}	737.6	1.341	238.9	1	1000
1 WATT =	3.413	44.25	0.7376	1.341×10^{-3}	0.2389	0.001	1

ELECTRIC CHARGE

	abcoul	amp-hr	COUL	faraday	statcoul
1 abcoulomb (1 emu) =	1	2.778×10^{-3}	10	1.036×10^{-4}	2.998×10^{10}
1 ampere-hour =	360	1	3600	3.730×10^{-2}	1.079×10^{13}
1 COULOMB =	0.1	2.778×10^{-4}	1	1.036×10^{-5}	2.998×10^{9}
1 faraday =	9652	26.81	9.652×10^{4}	1	2.893×10^{14}
1 statcoulomb (1 esu) =	3.336×10^{-11}	9.266×10^{-14}	3.336×10^{-10}	3.456×10^{-15}	1

1 electronic charge = 1.602×10^{-19} coulomb

ELECTRIC CURRENT

	abamp	AMP	statamp
1 abampere (1 emu) =	1	10	2.998×10^{10}
1 AMPERE =	0.1	1	2.998×10^{9}
1 statampere (1 esu) =	3.336×10^{-11}	3.336×10^{-10}	1

ELECTRIC POTENTIAL, ELECTROMOTIVE FORCE

	abv	VOLTS	statv
1 abvolt (1 emu) =	1	10^{-8}	3.336×10^{-11}
1 VOLT =	10^{8}	1	3.336×10^{-3}
1 statvolt (1 esu) =	2.998×10^{10}	299.8	1

ELECTRIC RESISTANCE

	abohm	OHMS	statohm
1 abohm (1 emu) =	1	10^{-9}	1.113×10^{-21}
1 OHM =	10^{9}	1	1.113×10^{-12}
1 statohm (1 esu) =	8.987×10^{20}	8.987×10^{11}	1

ELECTRIC RESISTIVITY

	abohm-cm	μohm-cm	ohm-cm	statohm-cm	OHM-M	ohm-circ mil/ft
1 abohm-centimeter (1 emu) =	1	0.001	10^{-9}	1.113×10^{-21}	10^{-11}	6.015×10^{-3}
1 micro-ohm-centimeter =	1000	1	10^{-6}	1.113×10^{-18}	10^{-8}	6.015
1 ohm-centimeter =	10^{9}	10^{6}	1	1.113×10^{-12}	0.01	6.015×10^{6}
1 statohm-centimeter (1 esu) =	8.987×10^{20}	8.987×10^{17}	8.987×10^{11}	1	8.987×10^{9}	5.406×10^{18}
1 OHM-METER =	10^{11}	10^{8}	100	1.113×10^{-10}	1	6.015×10^{8}
1 ohm-circular mil per foot =	166.2	0.1662	1.662×10^{-7}	1.850×10^{-19}	1.662×10^{-9}	1

CAPACITANCE

	abf	FARADS	μf[a]	statf
1 abfarad (1 emu) =	1	10^{9}	10^{15}	8.987×10^{20}
1 FARAD =	10^{-9}	1	10^{6}	8.987×10^{11}
1 microfarad =	10^{-15}	10^{-6}	1	8.987×10^{5}
1 statfarad (1 esu) =	1.113×10^{-21}	1.113×10^{-12}	1.113×10^{-6}	1

[a] This unit is frequently abbreviated as mf.

INDUCTANCE

	abhenry	HENRYS	μh	mh	stathenry
1 abhenry (1 emu) =	1	10^{-9}	0.001	10^{-6}	1.113×10^{-21}
1 HENRY =	10^{9}	1	10^{6}	1000	1.113×10^{-12}
1 microhenry =	1000	10^{-6}	1	0.001	1.113×10^{-18}
1 millihenry =	10^{6}	0.001	1000	1	1.113×10^{-15}
1 stathenry (1 esu) =	8.987×10^{20}	8.987×10^{11}	8.987×10^{17}	8.987×10^{14}	1

MAGNETIC FLUX

	maxwell	kiloline	WEBER
1 maxwell (1 line or 1 emu) =	1	0.001	10^{-8}
1 kiloline =	1000	1	10^{-5}
1 WEBER =	10^{8}	10^{5}	1

1 esu = 299.8 webers

MAGNETIC INDUCTION B

	gauss	kiloline/in.^{2}	WEBER/ METER^2 = TESLA	milligauss	
1 gauss (line per cm^2) =	1	6.452×10^{-3}	10^{-4}	1000	10^5
1 kiloline per in.^{2} =	155.0	1	1.550×10^{-2}	1.550×10^5	1.550×10^7
1 WEBER per METER^2 = 1 TESLA =	10^4	64.52	1	10^7	10^9
1 milligauss =	0.001	6.452×10^{-6}	10^{-7}	1	100
1 gamma =	10^{-5}	6.452×10^{-8}	10^{-9}	0.01	1

1 esu = 2.998×10^6 webers/meter^2

MAGNETOMOTIVE FORCE

	abamp-turn	AMP-TURN	gilbert
1 abampere-turn =	1	10	12.57
1 AMPERE-TURN =	0.1	1	1.257
1 gilbert =	7.958×10^{-2}	0.7958	1

1 pragilbert = 4π amp-turn   1 esu = 2.655×10^{-11} amp-turn

MAGNETIC FIELD STRENGTH H

	abamp-turn/cm	amp-turn/cm	amp-turn/in.	AMP-TURN/ METER	oersted
1 abampere-turn per centimeter =	1	10	25.40	1000	12.57
1 ampere-turn per centimeter =	0.1	1	2.540	100	1.257
1 ampere-turn per inch =	3.937×10^{-2}	0.3937	1	39.37	0.4947
1 AMPERE-TURN per METER =	0.001	0.01	2.540×10^{-2}	1	1.257×10^{-2}
1 oersted =	7.958×10^{-2}	0.7958	2.021	79.58	1

1 oersted = 1 gilbert   1 esu = 2.655×10^{-9} amp-turn/meter

1 praoersted = 4π amp-turn/meter

Mathematical Symbols and the Greek Alphabet

APPENDIX H

Mathematical signs and symbols

$=$ equals
$\cong$ equals approximately
$\neq$ is not equal to
$\equiv$ is identical to, is defined as
$>$ is greater than ($\gg$ is much greater than)
$<$ is less than ($\ll$ is much less than)
$\geqq$ is more than or equal to (or, is no less than)
$\leqq$ is less than or equal to (or, is no more than)
$\pm$ plus or minus (e.g., $\sqrt{4} = \pm 2$)
$\propto$ is proportional to (e.g., Hooke's law: $F \propto x$, or $F = -kx$)
Σ the sum of
$\bar{x}$ the average value of x

The Greek alphabet

Alpha	Α	α	Nu	Ν	ν
Beta	Β	β	Xi	Ξ	ξ
Gamma	Γ	γ	Omicron	Ο	ο
Delta	Δ	δ	Pi	Π	π
Epsilon	Ε	ϵ	Rho	Ρ	ρ
Zeta	Ζ	ζ	Sigma	Σ	σ
Eta	Η	η	Tau	Τ	τ
Theta	Θ	θ, ϑ	Upsilon	ϒ	υ
Iota	Ι	ι	Phi	Φ	ϕ, φ
Kappa	Κ	κ	Chi	Χ	χ
Lambda	Λ	λ	Psi	Ψ	ψ
Mu	Μ	μ	Omega	Ω	ω

Mathematical Formulas

APPENDIX I

Quadratic formula

If $ax^2 + bx + c = 0$, then $x = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$.

Trigonometric functions of angle θ

$$\sin\theta = \frac{y}{r} \qquad \cos\theta = \frac{x}{r}$$

$$\tan\theta = \frac{y}{x} \qquad \cot\theta = \frac{x}{y}$$

$$\sec\theta = \frac{r}{x} \qquad \csc\theta = \frac{r}{y}$$

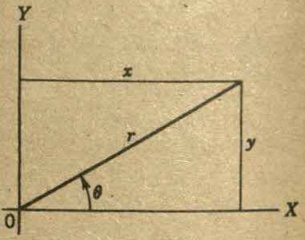


Fig. App. I

Pythagorean theorem

$$x^2 + y^2 = r^2$$

Trigonometric identities

$$\sin^2\theta + \cos^2\theta = 1 \qquad \sec^2\theta - \tan^2\theta = 1 \qquad \csc^2\theta - \cot^2\theta = 1$$

$$\sin(\alpha \pm \beta) = \sin\alpha\cos\beta \pm \cos\alpha\sin\beta$$

$$\cos(\alpha \pm \beta) = \cos\alpha\cos\beta \mp \sin\alpha\sin\beta$$

$$\tan(\alpha \pm \beta) = \frac{\tan\alpha \pm \tan\beta}{1 \mp \tan\alpha\tan\beta}$$

$$\sin 2\theta = 2\sin\theta\cos\theta$$

$$\cos 2\theta = \cos^2\theta - \sin^2\theta = 2\cos^2\theta - 1 = 1 - 2\sin^2\theta$$

$$\sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i} \qquad \cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2}$$

$$e^{\pm i\theta} = \cos\theta \pm i\sin\theta$$

Taylor's series

$$f(x_0 + x) = f(x_0) + f'(x_0)x + f''(x_0)\frac{x^2}{2!} + f'''(x_0)\frac{x^3}{3!} + \cdots$$

Series expansions (these expansions converge for $-1 < x < 1$, except as noted)

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \cdots$$

$$\sqrt{1+x} = 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} - \cdots$$

$$\frac{1}{\sqrt{1+x}} = 1 - \frac{x}{2} + \frac{3x^2}{8} - \frac{5x^3}{16} + \cdots$$

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots \qquad (-\infty < x < \infty)$$

$$x \text{ in radians} \begin{cases} \sin x = x - \dfrac{x^3}{6} + \dfrac{x^5}{120} - \cdots & (-\infty < x < \infty) \\ \cos x = 1 - \dfrac{x^2}{2} + \dfrac{x^4}{24} - \cdots & (-\infty < x < \infty) \\ \tan x = x + \dfrac{x^3}{3} + \dfrac{2x^5}{15} + \cdots & (-\pi/2 < x < \pi/2) \end{cases}$$

$$(x+y)^n = x^n + \frac{n}{1!}x^{n-1}y + \frac{n(n-1)}{2!}x^{n-2}y^2 + \cdots \qquad (x^2 > y^2)$$

Derivatives and indefinite integrals

In what follows, the letters u and v stand for any functions of x, and a and m are constants. To each of the integrals should be added an arbitrary constant of integration. *A Short Table of Integrals* by Peirce and Foster (Ginn and Co.) gives a more extensive tabulation.

1. $\dfrac{dx}{dx} = 1$
2. $\dfrac{d}{dx}(au) = a\dfrac{du}{dx}$
3. $\dfrac{d}{dx}(u + v) = \dfrac{du}{dx} + \dfrac{dv}{dx}$
4. $\dfrac{d}{dx}x^m = mx^{m-1}$
5. $\dfrac{d}{dx}\ln x = \dfrac{1}{x}$
6. $\dfrac{d}{dx}(uv) = u\dfrac{dv}{dx} + v\dfrac{du}{dx}$
7. $\dfrac{d}{dx}e^x = e^x$
8. $\dfrac{d}{dx}\sinh x = \cosh x$

1. $\int dx = x$
2. $\int au\,dx = a\int u\,dx$
3. $\int (u + v)\,dx = \int u\,dx + \int v\,dx$
4. $\displaystyle\int x^m\,dx = \frac{x^{m+1}}{m+1} \qquad (m \neq -1)$
5. $\displaystyle\int \frac{dx}{x} = \ln |x|$
6. $\displaystyle\int u\frac{dv}{dx}\,dx = uv - \int v\frac{du}{dx}\,dx$
7. $\int e^x\,dx = e^x$
8. $\int \cosh x\,dx = \sinh x$

9. $\frac{d}{dx}\cosh x = \sinh x$

10. $\frac{d}{dx}\arctan x = \frac{1}{1+x^2}$

11. $\frac{d}{dx}\arcsin x = \frac{1}{\sqrt{1-x^2}}$

12. $\frac{d}{dx}\operatorname{arcsec} x = \frac{1}{x\sqrt{x^2-1}}$

13. $\frac{d}{dx}\cos x = -\sin x$

14. $\frac{d}{dx}\sin x = \cos x$

15. $\frac{d}{dx}\tan x = \sec^2 x$

16. $\frac{d}{dx}\cot x = -\csc^2 x$

17. $\frac{d}{dx}\sec x = \tan x \sec x$

18. $\frac{d}{dx}\csc x = -\cot x \csc x$

9. $\int \sinh x\,dx = \cosh x$

10. $\int \frac{dx}{1+x^2} = \arctan x$

11. $\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x$

12. $\int \frac{dx}{x\sqrt{x^2-1}} = \operatorname{arcsec} x$

13. $\int \sin x\,dx = -\cos x$

14. $\int \cos x\,dx = \sin x$

19. $\int \tan x\,dx = \ln|\sec x|$

20. $\int \cot x\,dx = \ln|\sin x|$

21. $\int \sec x\,dx = \ln|\sec x + \tan x|$

22. $\int \csc x\,dx = \ln|\csc x - \cot x|$

Vector products

Let $\mathbf{i}, \mathbf{j}, \mathbf{k}$ be unit vectors in the x, y, z directions. Then

$$\mathbf{i}\cdot\mathbf{i} = \mathbf{j}\cdot\mathbf{j} = \mathbf{k}\cdot\mathbf{k} = 1, \qquad \mathbf{i}\cdot\mathbf{j} = \mathbf{j}\cdot\mathbf{k} = \mathbf{k}\cdot\mathbf{i} = 0,$$

$$\mathbf{i}\times\mathbf{i} = \mathbf{j}\times\mathbf{j} = \mathbf{k}\times\mathbf{k} = 0,$$

$$\mathbf{i}\times\mathbf{j} = \mathbf{k}, \qquad \mathbf{j}\times\mathbf{k} = \mathbf{i}, \qquad \mathbf{k}\times\mathbf{i} = \mathbf{j}.$$

Any vector $\mathbf{a}$ with components a_x, a_y, a_z along the x, y, z axes can be written

$$\mathbf{a} = a_x\mathbf{i} + a_y\mathbf{j} + a_z\mathbf{k}.$$

Let $\mathbf{a}, \mathbf{b}, \mathbf{c}$ be arbitrary vectors with magnitudes a, b, c. Then

$$\mathbf{a}\times(\mathbf{b}+\mathbf{c}) = \mathbf{a}\times\mathbf{b} + \mathbf{a}\times\mathbf{c}$$

$$(s\mathbf{a})\times\mathbf{b} = \mathbf{a}\times(s\mathbf{b}) = s(\mathbf{a}\times\mathbf{b}) \qquad (s \text{ a scalar}).$$

Let θ be the smaller of the two angles between $\mathbf{a}$ and $\mathbf{b}$. Then

$$\mathbf{a}\cdot\mathbf{b} = \mathbf{b}\cdot\mathbf{a} = a_xb_x + a_yb_y + a_zb_z = ab\cos\theta$$

$$\mathbf{a}\times\mathbf{b} = -\mathbf{b}\times\mathbf{a} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix} = (a_yb_z - b_ya_z)\mathbf{i} + (a_zb_x - b_za_x)\mathbf{j} + (a_xb_y - b_xa_y)\mathbf{k}$$

$$|\mathbf{a}\times\mathbf{b}| = ab\sin\theta$$

$$\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \mathbf{b}\cdot(\mathbf{c}\times\mathbf{a}) = \mathbf{c}\cdot(\mathbf{a}\times\mathbf{b})$$

$$\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\cdot\mathbf{c})\mathbf{b} - (\mathbf{a}\cdot\mathbf{b})\mathbf{c}$$

Values of Trigonometric Functions

APPENDIX J

TRIGONOMETRIC FUNCTIONS

Radians	Degrees	Sines	Cosines	Tangents	Cotangents		
0000	0	.0000	1.0000	.0000	∞	90	1.5708
.0175	1	.0175	.9998	.0175	57.29	89	1.5533
.0349	2	.0349	.9994	0349	28.64	88	1.5359
.0524	3	.0523	.9986	.0524	19.08	87	1.5184
.0698	4	.0698	.9976	.0699	14.30	86	1.5010
.0873	5	.0872	.9962	.0875	11.430	85	1.4835
.1047	6	1045	.9945	.1051	9.514	84	1.4661
.1222	7	.1219	.9925	.1228	8.144	83	1.4486
.1396	8	.1392	.9903	.1405	7.115	82	1.4312
.1571	9	.1564	.9877	.1584	6.314	81	1.4137
.1745	10	.1736	.9848	.1763	5.671	80	1.3963
.1920	11	.1908	.9816	.1944	5.145	79	1.3788
.2094	12	.2079	.9781	.2126	4.705	78	1.3614
2269	13	.2250	.9744	.2309	4.332	77	1.3439
.2443	14	.2419	.9703	.2493	4.011	76	1.3265
.2618	15	.2588	.9659	.2679	3.732	75	1.3090
.2793	16	.2756	.9613	.2867	3.487	74	1.2915
2967	17	.2924	.9563	.3057	3.271	73	1.2741
.3142	18	.3090	.9511	.3249	3.078	72	1.2566
.3316	19	.3256	.9455	.3443	2.904	71	1.2392
		Cosines	Sines	Cotangents	Tangents	Degrees	Radians

TRIGONOMETRIC FUNCTIONS (*Continued*)

Radians	Degrees	Sines	Cosines	Tangents	Cotangents		
.3491	20	.3420	.9397	.3640	2.748	70	1.2217
.3665	21	.3584	.9336	.3839	2.605	69	1.2043
.3840	22	.3746	.9272	.4040	2.475	68	1.1868
.4014	23	.3907	.9205	.4245	2.356	67	1.1694
.4189	24	.4067	.9135	.4452	2.246	66	1.1519
.4363	25	.4226	.9063	.4663	2.144	65	1.1345
.4538	26	.4384	.8988	.4877	2.050	64	1.1170
.4712	27	.4540	.8910	.5095	1.963	63	1.0996
.4887	28	.4695	.8829	.5317	1.881	62	1.0821
.5061	29	.4848	.8746	.5543	1.804	61	1.0647
.5236	30	.5000	.8660	.5774	1.732	60	1.0472
.5411	31	.5150	.8572	.6009	1.664	59	1.0297
.5585	32	.5299	.8480	.6249	1.600	58	1.0123
.5760	33	.5446	.8387	.6494	1.540	57	0.9948
.5934	34	.5592	.8290	.6745	1.483	56	0.9774
.6109	35	.5736	.8192	.7002	1.428	55	0.9599
.6283	36	.5878	.8090	.7265	1.376	54	0.9425
.6458	37	.6018	.7986	.7536	1.327	53	0.9250
.6632	38	.6157	.7880	.7813	1.280	52	0.9076
.6807	39	.6293	.7771	.8098	1.235	51	0.8901
.6981	40	.6428	.7660	.8391	1.192	50	0.8727
.7156	41	.6561	.7547	.8693	1.150	49	0.8552
.7330	42	.6691	.7431	.9004	1.111	48	0.8378
.7505	43	.6820	.7314	.9325	1.072	47	0.8203
.7679	44	.6947	.7193	.9657	1.036	46	0.8029
.7854	45	.7071	.7071	1.0000	1.000	45	0.7854
		Cosines	Sines	Cotangents	Tangents	Degrees	Radians

Nobel Prize Winners in Physics[a]

APPENDIX K

1901	Wilhelm Konrad Röntgen	1845–1923	German	Discovery of X-rays
1902	Hendrik Antoon Lorentz	1853–1928	Dutch	Influence of magnetism on the phenomena of atomic radiation.
	Pieter Zeeman	1865–1943	Dutch	
1903	Henri Becquerel	1852–1908	French	Discovery of natural radioactivity and of the radioactive elements radium and polonium.
	Pierre Curie	1850–1906	French	
	Marie Curie	1867–1934	French	
1904	Baron Rayleigh	1842–1919	English	Discovery of argon.
1905	Philipp Lenard	1862–1947	German	Research in cathode rays.
1906	Sir Joseph John Thomson	1856–1940	English	Conduction of electricity through gases.
1907	Albert A. Michelson	1852–1931	U. S.	Invention of interferometer and spectroscopic and metrological investigations.
1908	Gabriel Lippmann	1845–1921	French	Photographic reproduction of colors.
1909	Guglielmo Marconi	1874–1937	Italian	Development of wireless telegraphy.
	Karl Ferdinand Braun	1850–1918	German	
1910	Johannes Diderik van der Waals	1837–1923	Dutch	Equations of state of gases and fluids.
1911	Wilhelm Wien	1864–1928	German	Laws of heat radiation.
1912	Nils Gustaf Dalen	1869–1937	Swedish	Automatic coastal lighting.
1913	Heike Kamerlingh-Onnes	1853–1926	Dutch	Properties of matter at low temperatures; production of liquid helium.
1914	Max von Laue	1879–1960	German	Diffraction of X-rays in crystals.
1915	Sir William Henry Bragg	1862–1942	English	Study of crystal structure by means of X-rays.
	Sir William Lawrence Bragg	1890–	English—his son	
1916	(No award)			
1917	Charles Glover Barkla	1877–1944	English	Discovery of the characteristic X-rays of elements.
1918	Max Planck	1858–1947	German	Discovery of the elemental quantum.
1919	Johannes Stark	1874–1957	German	Discovery of the Doppler effect in canal rays and the splitting of spectral lines in the electric field.

1920	Charles Edouard Guillaume	1861–1938	Swiss	Discovery of the anomalies of nickel-steel alloys.
1921	Albert Einstein	1879–1955	German	Discovery of the law of the photoelectric effect.
1922	Niels Bohr	1885–1963	Danish	Study of structure and radiations of atoms.
1923	Robert Andrews Millikan	1868–1953	U. S.	Work on elementary electric charge and the photoelectric effect.
1924	Manne Siegbahn	1886–	Swedish	Discoveries in the area of X-ray spectra.
1925	James Franck	1882–1964	German	Laws governing collision between electron and atom.
	Gustav Hertz	1887	German	
1926	Jean Perrin	1870–1942	French	Discovery of the equilibrium of sedimentation.
1927	Arthur H. Compton	1892–1962	U. S.	Discovery of the scattering of X-rays by charged particles
	Charles T. R. Wilson	1869–1959	English	Invention of the cloud chamber, a device to make visible the paths of charged particles.
1928	Sir Owen Williams Richardson	1879–1959	English	Discovery of the law known by his name (the dependency of the emission of electrons on temperature).
1929	Louis-Victor de Broglie	1892–	French	Wave nature of electrons.
1930	Sir Chandrasekhara Raman	1888–	Indian	Work on the scattering of light and discovery of the effect known by his name.
1931	(No award)			
1932	Werner Heisenberg	1901–	German	Creation of quantum mechanics.
1933	Paul Adrien Maurice Dirac	1902–	English	Discovery of new fertile forms of the atomic theory.
	Erwin Schrödinger	1887–1961	Austrian	
1934	(No award)			
1935	James Chadwick	1891–	English	Discovery of the neutron.
1936	Victor Hess	1883–1964	Austrian	Discovery of cosmic radiation.
	Carl David Anderson	1905–	U. S.	Discovery of the positron.
1937	Clinton Joseph Davisson	1881–1958	U. S.	Discovery of diffraction of electrons by crystals.
	George P. Thomson	1892–	English	
1938	Enrico Fermi	1901–1954	Italian	Artificial radioactive elements from neutron irradiation.
1939	E. O. Lawrence	1901–1958	U. S.	Invention of the cyclotron.
1940–1942	(No awards)			
1943	Otto Stern	1888	U. S.[a]	Work with molecular beams and magnetic moment of proton.
1944	Isidor Isaac Rabi	1898–	U. S.	Nuclear magnetic resonance.
1945	Wolfgang Pauli	1900–1958	Austrian	Discovery of quantum exclusion principle.
1946	Percy Williams Bridgman	1882–1961	U. S.	High-pressure physics.
1947	Sir Edward Appleton	1892–	English	Upper atmosphere physics and discovery of Appleton layer.
1948	Patrick Maynard Stuart Blackett	1897–	English	Discoveries in cosmic radiation and nuclear physics.
1949	Hideki Yukawa	1907–	Japanese	Prediction of existence of meson.
1950	Cecil Frank Powell	1903–	English	Photographic method of studying nuclear processes; discoveries about mesons.
1951	Sir John Douglas Cockcroft	1897–	English	Transmutation of atomic nuclei by artificially accelerated atomic particles.
	Ernest Thomas Sinton Walton	1903–	Irish	
1952	Felix Bloch	1905–	U. S.	Measure of magnetic fields in atomic nuclei.
	Edward Mills Purcell	1912–	U. S.	
1953	Frits Zernike	1888–	Dutch	Invention of phase contrast microscopy.
1954	Max Born	1882–	English[b]	Work in quantum mechanics and statistical interpretation of wave function.

	Walther Bothe	1891–1957	German	Analysis of cosmic radiation using the coincidence method.
1955	Willis E. Lamb, Jr.	1913	U. S.	Fine structure of hydrogen.
	Polykarp Kusch	1911–	U. S.	Magnetic moment of electron.
1956	John Bardeen	1908–	U. S.	Invention and development of transistor.
	Walter H. Brattain	1902–	U. S.	
	William B. Shockley	1910–	U. S.[c]	
1957	Chen Ning Yang	1922–	Chinese[d]	Non-conservation of parity and work in elementary particle theory.
	Tsung Dao Lee	1926–	Chinese[d]	
1958	Pavel A. Čerenkov	1904–	Russian	Discovery and interpretation of Čerenkov effect of radiation by fast charged particles in matter.
	Ilya M. Frank	1908–	Russian	
	Igor Y. Tamm	1895–	Russian	
1959	Owen Chamberlain	1920–	U. S.	Discovery of the antiproton.
	Emilio Gino Segré	1905–	U. S.[e]	
1960	Donald A. Glaser	1926–	U. S.	Invention of bubble chamber.
1961	Robert L. Hofstadter	1915–	U. S.	Electromagnetic structure of nucleons from high-energy electron scattering.
	Rudolf L. Mössbauer	1929	German	Discovery of recoilless resonance absorption of gamma rays in nuclei.
1962	Len D. Landau	1908–	Russian	Theory of condensed matter; phenomena of superfluidity and superconductivity.
1963	Eugene B. Wigner	1902–	U. S.[f]	Contributions to theoretical atomic and nuclear physics.
	Maria Goeppert-Mayer	1906–	U. S.[g]	Shell model theory and magic numbers for the atomic nucleus.
	J. H. D. Jensen	1907–	German	
1964	C. H. Townes	1915–	U. S.	Invention of the maser and theory of coherent atomic radiation.
	Nikolai Basov	1922–	Russian	
	Aleksandr Prokhorov	1916–	Russian	
1965	Richard Feynman	1918	U. S.	Development of quantum electrodynamics
	Julian Schwinger	1918	U. S.	
	Shin-Ichiro Tomonaga	1906	Japanese	

---

[a] See *Nobel: The Man and His Prizes*, by Schück et al., Elsevier, N. Y.

[b] Born in Germany; naturalized British citizen.

[c] Born in England; naturalized U. S. citizen.

[d] Both have permanent U. S. resident status.

[e] Born in Italy; naturalized U. S. citizen.

[f] Born in Hungary; naturalized U. S. citizen.

[g] Born in Germany; naturalized U. S. citizen.

The Gaussian System of Units

APPENDIX L

Much of the literature of physics is written, and continues to be written, in the Gaussian system of units. In electromagnetism many equations have slightly different forms depending on whether it is intended, as in this book, that mks variables be used or that Gaussian variables be used. Equations in this book can be cast in Gaussian form by replacing the symbols listed below under "rationalized mks" by those listed under "Gaussian." For example, Eq. 37–26,

$$\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M})$$

becomes

$$\frac{\mathbf{B}}{c} = \left(\frac{4\pi}{c^2}\right)\left(\frac{c}{4\pi}\,\mathbf{H} + c\mathbf{M}\right)$$

or

$$\mathbf{B} = \mathbf{H} + 4\pi\mathbf{M}$$

in Gaussian form. Symbols used in this book that are not listed below remain unchanged. The quantity c is the speed of light.

Quantity	Rationalized mks	Gaussian
Permittivity constant	ϵ_0	$1/4\pi$
Permeability constant	μ_0	$4\pi/c^2$
Electric displacement	$\mathbf{D}$	$\mathbf{D}/4\pi$
Magnetic induction	$\mathbf{B}$	$\mathbf{B}/c$
Magnetic flux	Φ_B	Φ_B/c
Magnetic field strength	$\mathbf{H}$	$c\mathbf{H}/4\pi$
Magnetization	$\mathbf{M}$	$c\mathbf{M}$
Magnetic dipole moment	$\boldsymbol{\mu}$	$c\boldsymbol{\mu}$

In addition to casting the equations in the proper form it is of course necessary to use a consistent set of units in those equations. Below we list some equivalent quantities in mks and Gaussian units. This table can be used to transform units from one system to the other.

Quantity	Symbol	Mks system	Gaussian system
Length	l	1 meter	10^2 cm
Mass	m	1 kg	10^3 gm
Time	t	1 sec	1 sec
Force	$\mathbf{F}$	1 newton	10^5 dynes
Work or Energy	W, E	1 joule	10^7 ergs
Power	P	1 watt	10^7 ergs/sec
Charge	q	1 coulomb	3×10^9 statcoul
Current	i	1 ampere	3×10^9 statamp
Electric field strength	$\mathbf{E}$	1 volt/meter	$\frac{1}{3} \times 10^{-4}$ statvolt/cm
Electric potential	V	1 volt	$\frac{1}{300}$ statvolt
Electric polarization	$\mathbf{P}$	1 coul/meter^2	3×10^5 statcoul/cm^2
Electric displacement	$\mathbf{D}$	1 coul/meter^2	$12\pi \times 10^5$ statvolt/cm
Resistance	R	1 ohm	$\frac{1}{9} \times 10^{-11}$ sec cm^{-1}
Capacitance	C	1 farad	9×10^{11} cm
Magnetic flux	Φ_B	1 weber	10^8 maxwells
Magnetic induction	$\mathbf{B}$	1 tesla $\equiv$ 1 weber/meter^2	10^4 gauss
Magnetic field strength	$\mathbf{H}$	1 amp-turn/meter	$4\pi \times 10^{-3}$ oersted
Magnetization	$\mathbf{M}$	1 weber/meter^2	$1/4\pi \times 10^4$ gauss
Inductance	I	1 henry	$\frac{1}{9} \times 10^{-11}$

All factors of 3 in the above table, apart from exponents, should be replaced by (2.997925 ± 0.000003) for accurate work; this arises from the numerical value of the speed of light. For example the mks unit of capacitance (= 1 farad) is actually 8.98758×10^{11} cm rather than 9 ($= 3^2$) $\times 10^{11}$ cm as listed above. This example also shows that not only units but also the dimensions of physical quantities may differ between the two systems. In the mks system (see Appendix F) the dimensions of capacitance are $M^{-1}L^{-2}T^2Q^2$; in the Gaussian system they are simply L, the Gaussian standard unit of capacitance being 1 cm.

The student should consult *Classical Electromagnetism*, p. 611, by J. D. Jackson (John Wiley and Sons, 1962) for a fuller treatment of units and dimensions.

Answers to Odd-Numbered Problems

Chapter 26

1. 0.12 amp.

3. $\pm 2.4 \times 10^{-8}$ coul.

5. 6.0×10^{-8} coul.

7. (*a*) $Q = -2\sqrt{2}q$,
(*b*) no.

9. (*a*) 5.7×10^{13} coul,
(*b*) no,
(*c*) 660 tons.

11. 3.8×10^{-5} coul, 1.2×10^{-5} coul.

13. (*b*) Along the body diagonal.

15. (*a*) 6.3×10^{11},
(*b*) 7.3×10^{-13}.

17. (*a*) 510 nt,
(*b*) 7.7×10^{28} meters/sec^2.

Chapter 27

5. 5.6×10^{-11} coul.

9. (*a*) The larger charge produces a field of 13×10^4 nt/coul at the site of the small the smaller charge produces a field of 5.3×10^4 nt/coul at the site of the larger.
(*b*) 1.1×10^{-2} nt, repulsive.

11. **E** lies in the median plane and points radially away from the charge axis.

13. 1.0×10^5 nt/coul, pointing up.

17. $E = q/8\pi\epsilon_0 a^2$, pointing along the axis of symmetry and away from the hemisphe

21. (*a*) $E = \dfrac{1}{4\pi\epsilon_0}\dfrac{qx}{(a^2 + x^2)^{3/2}}$,
(*b*) $E = \dfrac{1}{2\pi^2\epsilon_0}\dfrac{(q_1 - q_2)a}{(a^2 + x^2)^{3/2}}$.

25. (*a*) 1.5×10^3 nt/coul,
(*b*) 2.4×10^{-16} nt (up),
(*c*) 1.6×10^{-26} nt,
(*d*) 1.5×10^{10}.

27. (*a*) 7.1 cm,
(*b*) 2.9×10^{-8} sec,
(*c*) 11%.

29. $\dfrac{dE}{dz} = -\dfrac{8q}{\pi\epsilon_0 l^3}$, where l is the distance between the charges; yes.

31. 1.64×10^{-19} coul.

Chapter 28

1. $\pi R^2 E$ (an expected result).

3. $EA \cos \theta$.

7. (a) 4.0×10^6 nt/coul,
(b) $E = 0$.

9. (a) $E = \sigma/\epsilon_0$, to the left,
(b) $E = 0$,
(c) $E = \sigma/\epsilon_0$, to the right.

11. 4.9×10^{-10} coul.

13. 0.44 mm.

19. (a) $E = \dfrac{q}{2\pi\epsilon_0 l r}$ (radially inward), where l is the length of the cylinders,
(b) $-q$ on inner surface and $-q$ on outer surface,
(c) $E = \dfrac{q}{2\pi\epsilon_0 l r}$ (radially outward).

21. 2.5×10^{-9} coul/meter^2.

23. 1.9×10^2 nt,
2.9×10^{28} meters/sec^2.

25. (a) 1.1 nt-meters^2/coul,
(b) 9.3×10^{-12} coul.

Chapter 29

1. 0.89 mm.

3. 900 volts.

5. (a) Between the charges a distance of 25 cm from $+q$ and outside the charges a distance of 50 cm from $+q$.
(b) Outside the charges a distance of 140 cm from $+q$.

7. $V_A - V_B = \dfrac{q}{2\pi\epsilon_0} \dfrac{d}{a(a + d)}$.

9. 1.9×10^{-29} coul meter; text value (0.61×10^{-29} coul meter) is lower and correct; the assumptions made in this problem are oversimplified.

11. 99 tons.

13. -6.4×10^{-7} joule.

15. (a) -0.12 volt,
(b) 1.8×10^{-8} nt/coul, radially inward.

17. 2900 volts.

21. No.

23. (a) -180 volts,
(b) $+2900$ volts and -9000 volts.

27. $-0.21q^2/\epsilon_0 a$.

29. (a) 1.1×10^{17} volts/meter,
(b) 4.6×10^{21} volts/meter, assuming a nuclear radius of 5×10^{-15} meter.

31. (a) 2.6×10^5 volts,
(b) $\dfrac{\sqrt{5}}{3}\,c = 0.745\,c$.

33. 9.0 kw.

35. (a) 3.2×10^{-13} joule,
(b) 1.6×10^{-13} joule,
(c) proton.

Chapter 30

1. (a) $q_2 = q_8 = 4.8 \times 10^{-4}$ coul,
$V_2 = 240$ volts,
$V_8 = 60$ volts;
(b) $q_2 = 2.0 \times 10^{-4}$ coul,
$q_8 = 7.7 \times 10^{-4}$ coul,
$V_2 = V_8 = 96$ volts;
(c) $q_2 = q_8 =$ zero,
$V_2 = V_8 =$ zero.

3. 43 $\mu\mu$f.

11. 3.2 μf.

13. 4 μf.

15. 7.3 μf.

21. Mica.

23. Assuming $\kappa = 5.4$:
(a) 10^4 volts/meter,
(b) $+5.0 \times 10^{-9}$ coul, on the positive plate,

(c) -4.1×10^{-9} coul, next to the positive plate.

25. 0.63 meter^2.

27. 0.11 joule/meter^3.

31. (a) 1.3×10^{-7} joule,
(b) no.

33. 7.0×10^{-6} coul.

35. 7.0¢.

37. (a) $q_1 = q_2 = 3.3 \times 10^{-4}$ coul,
$q_3 = 4.0 \times 10^{-4}$ coul;
(b) $V_1 = 33$ volts,
$V_2 = 67$ volts,
$V_3 = 100$ volts;
(c) $U_1 = 5.4 \times 10^{-3}$ joule,
$U_2 = 10.9 \times 10^{-3}$ joule,
$U_3 = 2.0 \times 10^{-2}$ joule.

Chapter 31

1. (a) 1200 coul,
(b) 7.5×10^{21} electrons.

3. (a) 2.4, iron being larger;
(b) no.

5. 6.7×10^{-6} coul/meter^2.

7. 54 ohms.

9. (a) 2.2×10^{-7} ohm,
(b) nickel ($\rho = 6.8 \times 10^{-8}$ ohm-meter).

11. (a) 260°C,
(b) yes.

13. 0.39% (ρ),
0.0017% (l),
0.0034% (A).

17. 11 ohms.

19. (a) 4.9×10^6 amp/meter^2,
(b) 8.4×10^{-2} volt/meter,
(c) 26 volts,
(d) 640 watts.

21. 620 watts.

23. (a) 8.7%,
(b) smaller.

Chapter 32

1. 1.1×10^4 joules.

3. -0.62%.

7. (a) 990 ohms,
(b) 9.9×10^{-4} watt.

9. $V_d - V_c = 1.3$ volt.

11. (a) 120 ohms,
(b) $i_1 = 0.051$ amp,
$i_2 = i_3 = 0.019$ amp,
$i_4 = 0.013$ amp.

13. (a) $R = \frac{1}{2}r$,
(b) $E^2/2r$.

17. 38 ohms.

19. (a) 10 ohms,
(b) 14 ohms,
(c) 10 ohms.

21. series: $i = \dfrac{2}{R + 2r}$.
parallel: $i = \dfrac{2}{2R + r}$.

(a) Put R_1 roughly in the middle of its range; adjust current roughly with R_2; make fine adjustment with R_1.

(b) Relatively large percentage changes in R_1 cause only small percentage changes in the resistance of the parallel combination, thus permitting fine adjustment. The ratio is 1:21.

27. 2.7%.

29. 4.6 time constants.

33. (a) 9.6×10^{-7}
(b) 1.1×10^{-6} watt,
(c) 2.7×10^{-6} watt,
(d) 3.8×10^{-6} watt.

Chapter 33

1. (*a*) East,
(*b*) 6.3×10^{14} meters/sec^2,
(*c*) 3.0 mm.

3. 7.5 nt, perpendicular to the wire and to **B**.

5. See Appendix.

7. 3.8 coul.

9. 520 gauss, normal to plane of tracks.

11. 4.3×10^{-3} nt-meter. The torque vector is parallel to the long side of the coil and points down.

15. (*a*) 1.4×10^{-4} meter/sec,
(*b*) 4.5×10^{-23} nt (down),
(*c*) 2.8×10^{-4} volt/meter (down),
(*d*) 5.7×10^{-6} volt (top +, bottom −),
(*e*) same as (*b*).

17. (*a*) $K_p = K_d = \frac{1}{2}K_\alpha$,
(*b*) $R_d = 14$ cm.
(*c*) $R_\alpha = 14$ cm.

19. (*a*) 2.6×10^7 meters/sec,
(*b*) 1.1×10^{-7} sec,
(*c*) 14 Mev,
(*d*) 7.0×10^6 volts.

23. 1.6×10^{-8} weber/meter^2, horizontal and at right angles to the equator.

25. $T = 3.6 \times 10^{-10}$ sec,
$p = 0.17$ mm,
$r = 1.5$ mm.

27. 2.11×10^{-25} kg or 127 proton masses.

29. (*a*) Increase,
(*b*) decrease.

31. (*a*) 8.5 Mev,
(*b*) 0.80 weber/meter^2,
(*c*) 34 Mev,
(*d*) 24 mc/sec
(*e*) 34 Mev, 1.6 webers/meter^2,
34 Mev, 12 mc/sec.

33. 1.4.

35. 3800 meters/sec.

Chapter 34

1. 7.9×10^{-3} weber/meter^2.

3. (*a*) 3.2×10^{-16} nt, parallel to current;
(*b*) 3.2×10^{-16} nt, radially outward if **v** is parallel to the current;
(*c*) zero.

7. 1.0×10^{-6} weber/meter.

9. **B** = 0 along a line parallel to the wire and 4.0 mm from it. If the current is horizontal and points toward the observer and the external field points horizontally from left to right, the line is directly above the wire.

11. 8.0×10^{-5} weber/meter^2, up.

13. 3.2×10^{-3} nt, toward the long wire.

23. (*a*) 1.0×10^{-3} weber/meter^2, out of figure;
(*b*) 8.0×10^{-4} weber/meter^2, out of figure.

27. (*a*) 9.4×10^{-5} weber/meter^2,
(*b*) 1.5×10^{-6} nt-meter.

Chapter 35

1. 2.0×10^{-2} coul.

3. Zero.

9. $V_m = \pi^2 BR^2 f$,
$i_m = \pi^2 BR^2 f/R_M$

11. 3.0×10^{-4} volt.

13. (*a*) 3.1×10^{-2} webers/sec,
(*b*) left to right.

17. (*a*) 4.4×10^7 meters/sec^2, to the right;
(*b*) zero;
(*c*) 4.4×10^7 meters/sec^2, to the left

Chapter 36

1. Let the current charge at 10 amp/sec.
9. 12 sec.
15. (*a*) 10 amp,
    (*b*) 100 joules.
21. (*a*) 2.5×10^{-6} joule/meter,
    (*b*) 14×10^{-6} joule/meter,
    (*c*) 0.8×10^{-6} joule/meter.
7. 1.0×10^{-7} weber.
13. 27 amp/sec.
17. 0.63 joule/meter^3.
23. 1.5×10^8 volt/meter.

Chapter 37

1. (*a*) 5.0×10^7 amp,
    (*b*) yes,
    (*c*) no.
5. (*a*) 7.6 amp meter^2,
    (*b*) 11 nt-meter.
11. (*a*) 1.8 webers/meter^2,
    (*b*) 6.5×10^{-23} joule.
3. (*a*) 1.4×10^{11} volts/meter,
    (*b*) 2.8×10^{-3} weber/meter^2.
9. 7.5×10^{-6} weber/meter^2.

Chapter 38

1. 600, 710, 1100, and 1300 cycles/sec.
15. $\omega = \sqrt{\dfrac{1}{L_1C_1}} = \sqrt{\dfrac{1}{L_2C_2}}$.
21. (*a*) 1.4×10^{14} volts/meter-sec,
    (*b*) 9.9×10^{-6} weber/meter^2.
13. (*a*) 35 cycles/sec.
    (*b*) 38 or 33 cycles/sec.
19. $i_d = \epsilon_0\pi r^2 \dfrac{dE}{dt} \qquad (r \leq R),$
    $i_d = \epsilon_0\pi R^2 \dfrac{dE}{dt} \qquad (r \geq R).$

Chapter 39

3. 4.9 cm, 5.2×10^8 meters/sec
    (= 1.7 c).
13. (*a*) a^2EB/μ_0
    for faces parallel to *xy*-plane, zero for others;
    (*b*) zero.
15. 1000 volts/meter,
    3.4×10^{-6} weber/meter^2.
9. 18 cm, 12 cm.

Chapter 40

1. (*a*) 5100 A and 6100 A,
    (*b*) 5.5×10^{14} cycles/sec and 1.8×10^{-15} sec.
3. $F_{\text{rad}} = 6.0 \times 10^8$ nt,
    $F_{\text{gr}} = 3.6 \times 10^{22}$ nt.
11. 0.13.
15. 3.8×10^{-2} A.
7. 1.3×10^{-7} nt/meter^2.
13. (*b*) 7.3 (cycles/sec)/(miles/hour).
17. Yellow-orange.

Chapter 41

3. 2.05×10^8 meters/sec.
15. (*b*) 0.17.
19. Cover the center of each face with a circle of radius 0.33 cm. The fraction covere
    is 0.35.
5. 1.56.

Chapter 42

1. (*a*) 7,
    (*b*) 5,
    (*c*) 2.
3. 40 cm.

6. Alternate vertical columns:
   (a) +, +40, -20, +2, no, yes;
   (c) concave, +40, +60, -2, yes, no;
   (e) convex, -20, +20, +0.5, no, yes;
   (g) -20, $-$, $-$, +5, +0.80, no, yes.
9. 4.24 cm.
11. Alternate vertical columns:
   (a) +, X, X, +20, X, -1, yes, no;
   (c) converging, +, X, X, -10, X, no, yes;
   (e) converging, +30, -15, +1.5, no, yes;
   (g) diverging, -120, -9.2, +0.92, no, yes;
   (i) converging, +3.3, X, X, +5, X, $-$, no.
17. 4.5 cm, 9.0 cm.

7. Object at center of curvature.
10. Alternate vertical columns:
   (a) -18, no;
   (c) +71, yes;
   (e) +30, no;
   (g) -26, no.

19. Assuming the light is incident from the left, a distance of $-\dfrac{r(n-2)}{2(n-1)^2}$ to the right of the right edge of the sphere.
21. (a) R' is negative and R'' is positive,
   (b) $i = \dfrac{-2r}{n+1}$,
   (c) virtual and erect.
23. (a) Coincides in location with original object and is enlarged 5.0 times,
   (c) virtual and inverted.

Chapter 43

1. Slit separation must be 0.034 mm.
5. 6.6×10^{-3} mm.
9. $I = \dfrac{I_m}{9}\left[1 + 8\cos^2\left(\dfrac{\pi d \sin\theta}{\lambda}\right)\right]$.
13. 80 million.
17. 4800 A (blue).
23. 5880 A.

3. 0.15°.
7. 3.0 mm.
11. $y = 17.4 \sin(\omega t + 13.3°)$.
15. 1.21.
19. 6700 A.
25. 6057.8021 A.

Chapter 44

1. 0.17 mm.
3. (a) $\lambda_a = 2\lambda_b$,
   (b) coincidences occur when $m_b = 2m_a$.
5. 79.7°
9. 9100 meters.
15. 3.

7. (a) 52.9°,
   (b) 10.2°,
   (c) 5.1°.
11. (a) 0.16 sec of arc,
   (b) 7.4×10^7 km,
   (c) 2.2×10^{-4} mm.
17. (a) Must have $d = 4a$,
   (b) Every fourth fringe is missing.

Chapter 45

1. Three complete orders.
5. The intensity would be concentrated near the twentieth order for blue and the eleventh order for red. The orders would overlap to such an extent as to appear almost white.

7. All wavelengths shorter than 6300 A.

9. 0°, ±10°, ±21°, ±32°, ±45°, and ±62°.

11. 5200 A to 6200 A.

17. (*a*) 0.0032°/A,
0.0077°/A,
0.024°/A;
(*b*) 40,000,
80,000,
120,000.

19. 3600 lines.

21. (*a*) 4.6×10^{-2} A.
(*b*) No. The resolution could normally be improved by going to a higher or diffraction, but in this case $m = 3$ is the highest order that can exist (assum that the light falls on the grating at right angles).

23. (*a*) 6×10^4 A,
(*b*) 1.5×10^4 A,
(*c*) 0, 1, 2, 3, 5, 6, 7, 9. The tenth order is at $\theta = 90°$.

25. 33°, 20°, 5.2° (all clockwise);
14° (counterclockwise).

27. Yes, $n = 3$ for $\lambda = 1.29$ A,
$n = 4$ for $\lambda = 0.97$ A.

Chapter 46

1. (*a*) ±55°,
(*b*) ±35°.

3. Assuming that a right-handed coordinate system is used,
(*a*) circular, counterclockwise as seen facing the source;
(*b*) elliptical, counterclockwise as seen facing the source, major axis of ellipse along $y = x$;
(*c*) plane, along $y = -x$.

5. 55°30′ to 55°46′.

11. 2.7×10^{-14} kg-meter^2/sec^2,
2.1 hr.

Chapter 47

1. 4.8 A.

3. 6.2×10^{23} photons/sec.

7. 5.9×10^{-6} ev.

9. (*a*) 2.0 ev,
(*b*) zero,
(*c*) 2.0 volts,
(*d*) 3000 A.

11. 100 years.

17. (*a*) 1,
(*b*) 5.3×10^{-11} meter,
(*c*) 1.1×10^{-34} joule-sec,
(*d*) 2.0×10^{-24} kg-meter/sec,
(*e*) 4.1×10^{15} radians/sec,
(*f*) 2.2×10^{6} meters/sec,
(*g*) 8.2×10^{-8} nt,
(*h*) 9.0×10^{22} meters/sec^2
(*i*) +13.6 ev,
(*j*) −27.2 ev,
(*k*) −13.6 ev.

19. +0.21 ev.

21. 2.6 ev.

25. +54 ev, assuming the electron to be in its ground state initially.

27. (*a*) 2.6×10^{-13} meter,
(*b*) 2800 ev,
(*c*) 4.4 A.

Chapter 48

1. (*a*) 1.7×10^{-35} meter.
3. (*a*) 3.3×10^{-24} kg-meter/sec, for each;
   (*b*) 38 ev for the electron and 6200 ev for the photon.
5. (*a*) Higher orders cannot exist for this accelerating potential and for these planes.
   (*b*) 59°; the crystal must be rotated with respect to the incident beam to satisfy Bragg's law for this new wavelength.
7. 1.5 A.
9. (*a*) 0.20,
   (*b*) 0.40,
   (*c*) 0.33,
   (*d*) 0.33.
11. 0.32.
15. 6.6×10^{-23} kg-meter/sec.

Answers to Odd-Numbered Supplementary Problems

Chapter 26

1. $-\frac{4}{9}q$, located $\frac{l}{3}$ from $+q$.

3. $\pm 1.0 \times 10^{-6}$ coul, $\mp 3.0 \times 10^{-6}$ coul.
5. 1.61×10^{-9} meter.

Chapter 27

1. $2\pi \sqrt{\dfrac{l}{g - \dfrac{q}{m}E}}$, $2\pi \sqrt{\dfrac{l}{g + \dfrac{q}{m}E}}$.

3. $\dfrac{Q}{\pi^2 \epsilon_0 R^2}$, vertically downward.

5. $2\pi \sqrt{\dfrac{pE}{I}}$.

Chapter 28

1. $+3.54 \times 10^{-6}$ coul.
3. 1.98×10^{7} meters/sec.

5. $A = \dfrac{Q}{2\pi a^2}$.

Chapter 29

1. 1.2×10^{-8} coul/meter^2.

3. (a) 1.0×10^{-8} coul from small to large sphere;
   (b) small sphere: 2×10^{-8} coul, 3000 volts,
   large sphere: 4×10^{-8} coul, 3000 volts.

9. $W = +\dfrac{Q}{8\pi\epsilon_0}\left[\dfrac{1}{r_1} - \dfrac{1}{r_2}\right]$.

11. (a) $\dfrac{qQ}{4\pi\epsilon_0 K}$.

    (b) $\sqrt{\dfrac{K}{m}}$.

13. (a) $\dfrac{\lambda}{4\pi\epsilon_0} \ln\left(\dfrac{y + L}{y}\right)$.

    (b) $\dfrac{-\lambda}{4\pi\epsilon_0} \dfrac{L}{y(y + L)}$.

    (c) Zero.

Chapter 30

5. $q_1 = \dfrac{C_1C_2 + C_1C_3}{C_1C_2 + C_1C_3 + C_2C_3} C_1V_0$

   $q_2 = q_3 =$

   $\dfrac{C_2C_3}{C_1C_2 + C_1C_3 + C_2C_3} C_1V_0$

7. (*a*) $10\epsilon_0$ farad,
   (*b*) $13.6\epsilon_0$ farad,
   (*c*) $1200\epsilon_0$ coul, $1200\epsilon_0$ coul,
   (*d*) 10^4 volts/meter,
   (*e*) 2.1×10^3 volts/meter,
   (*f*) 88 volts,
   (*g*) $19 \times 10^3\epsilon_0$ joules.

Chapter 31

1. (*a*) 2.3×10^{12},
   (*b*) 5.0×10^3,
   (*c*) 10^7 volts.
5. (*a*) $\rho_A = .10$ ohm-meter, $\rho_B = .05$ ohm-meter,
   (*b*) $E_A = 1.0$ volt/meter, $E_B = .50$ volts/meter,
   (*c*) $j_A = j_B = 10$ amp/meter²,
   (*d*) $V_A = 40$ volts, $V_B = 20$ volts.
7. $T = 130°$ C.

Chapter 32

1. $R = r_1 - r_2$.
3. (*a*) r,
   (*b*) R.
5. (*a*) $\frac{7}{12}R$,
   (*b*) $\frac{3}{4}R$,
   (*c*) $\frac{5}{6}R$.
7. (*a*) 85.0 ohm,
   (*b*) 85.0 ohm.
9. (*a*) 1.0×10^{-3} coul.,
   (*b*) 1.0×10^{-3} amp.,
   (*c*) $V_C = V_R = 1000e^{-t}$ volts,
   (*d*) e^{-2t} watt.

Chapter 33

1. 1 is positively charged, 2 is neutral, and 3 is negatively charged.
3. 2.7×10^{-4} weber/meter².
5. (*a*) 542 ohms, series,
   (*b*) 2.52 ohms, parallel.
7. (*a*) 2.8 Mc/sec.
   (*b*) 0.34 meter

Chapter 34

1. $-10\mu_0 i_0$.
3. $\dfrac{3\sqrt{2}\,\mu_0 i^2}{4\pi a}$, toward center of square.
5. zero.
7. $\dfrac{\mu_0 i}{4}\left(\dfrac{1}{R_1} - \dfrac{1}{R_2}\right)$, into page.
9. (*a*) $\dfrac{\mu_0 i}{2R}\left(1 + \dfrac{1}{\pi}\right)$, out of page.
   (*b*) $\dfrac{\mu_0 i}{2\pi R}\sqrt{1 + \pi^2}$, out of page at angle of 18° with page.

Chapter 35

3. (*a*) $\dfrac{\mu_0 i \pi r^2 R^2}{2x^3}$,
   (*b*) $\dfrac{3\mu_0 \pi i r^2 v}{2R^2 N^4}$.
   (*c*) In the same sense as the current in the larger loop.
5. (*a*) $q = \dfrac{-1}{R}\left[\Phi_B(t_2) - \Phi_B(t_1)\right]$,
   (*b*) No. B field could still have changed between t_1 and t_2 inducing currents.

Chapter 36

3. (*a*) $i_1 = 3.3$ amp, $i_2 = 3.3$ amp;
   (*b*) $i_1 = 4.5$ amp, $i_2 = 2.7$ amp;
   (*c*) $i_1 = 0$, $i_2 = 1.8$ amp;
   (*d*) $i_1 = 0$, $i_2 = 0$.
5. (*a*) meter, (*b*) weber, (*c*) watt, (*d*) coulomb, (*e*) ohm.

Chapter 37

1. $\dfrac{q\omega r^2}{2}$, in direction of ω.
3. (*a*) $(N \pm 1)\dfrac{eB}{m}$,
   (*b*) 744×10^{10} rad/sec; 759×10^{10} rad/sec.
5. 29 amp.

Chapter 38

1. Let T_1 be the period of inductor and 900 μf capacitor and T_2 the period of inductor and 100 μf capacitor. Then: (*a*) close S_2, wait $T_1/4$; (*b*) quickly close S_1 then open S_2; (*c*) wait $T_2/4$ and then open S_1.

Chapter 39

1. (*a*) $\dfrac{2\pi\epsilon_0}{\ln(b/a)}$
   (*b*) $\dfrac{\mu_0 \ln(b/a)}{2\pi}$
5. (*a*) 6.3×10^{-2} volt/meter,
   (*b*) 2.1×10^{-10} weber/meter²,
   (*c*) 1.3×10^4 watts.

7. (*a*) $B = \mu_0 \sigma R \omega$.
   (*b*) $E = \frac{1}{2}\mu_0 \sigma R^2 \alpha$.
   (*c*) $S = \frac{\mu_0}{2}\sigma^2 R^3 \omega \alpha$.
   (*d*) They are both $\mu_0 \pi \sigma^2 R^4 l \omega \alpha$.

Chapter 40

1. (*a*) 1.0×10^8 cycles/sec,
   (*b*) 1.0×10^{-6} weber/meter^2,
   (*c*) 2.1 meter^{-1}, 6.3×10^8 rad/sec,
   (*d*) 120 watts/meter^2,
   (*e*) 4.0×10^{-7} A nt, 4.0×10^{-7} nt/meter^2.
3. (*b*) About 6×10^{-7} meter (comparable with the wavelength of light).
   (*c*) No.

Chapter 41

1. 1.95×10^8 meters/sec.
3. 180 cm.
5. (*a*) $\sqrt{1 + \sin^2 \theta_1}$
   (*b*) $\sqrt{2}$

Chapter 42

1. 6.
5. (*b*) No.
   (*c*) Rays pass undeviated through system.

Chapter 43

1. .072 mm.
3. 8 microns.
7. 8400 A.
9. 1.89 microns.

Chapter 44

1. (*a*) The droplets act as diffracting obstacles producing the observed ring.
   (*b*) .20 mm.
3. 71 ft.

Chapter 45

3. (*a*) 5860 A.
   (*b*) Since the fourth maximum is missing, the slit width must lie between 1.00 and 1.69 microns.
5. (*a*) One line to each side of the central maximum.
   (*b*) 1.44×10^{-4} rad.
   (*c*) $\Delta\theta = \frac{1}{R}$

Chapter 46

1. 67% polarized, 33% unpolarized.
3. Side A has the Polaroid, and the Polaroid axis must be at 45° to the principal axes of the quarter-wave plate.

Index

Aberrations, for lenses, 1061
  for mirrors, 1048
Absorption spectrum, 1189
Acoustic cavity, 954, 977, 1203
Action-at-a-distance, 663, 652
Addition of velocities, 1006
Airy, Sir George, 852, 1112
Allen, A. J., 831
Alpha-emission, 658
Alpha-particle, 700
Alternating current generator, 895
Ammeter, 800
Ampere, 652
Ampere, A. M., 844, 853
Amperian current, 936
Ampere's law, 844, 847, 857, 911, 936, 961, 985
Amplitude grating, 1127
Analyzer, 1150
Andrade, E. N., 700, 701
Angstrom, 994
Angular momentum, of electrons, 916, 1193, 1204
  of light, 1167
Anisotropic, 1155
Annihilation, 660
Antiferromagnetism, 930
Aragò, D. F., 761, 1148
Arrington, W., 1136
Atomic number, 658

B, *see* Magnetic induction
Ballistic galvanometer, 928
Barkla, C. G., 1169
Bees and polarized light, 1169
Betatron, 884
Beth, 1167
B-H curve, 937
Biaxial crystal, 1162
Binding energy, 1195
Biot-Savart law, 859
Birefringence, 1156
Birge, R. T., 1002
Bloch, F., 931
Bloembergen, N., 934
Bohr, N., 1191, 1196, 1213
Bohr model, 864
Bol, 1102
Boltzmann, L., 963
Boltzmann's constant, 1179
Born, M., 1208
Bradley, J., 999
Boundary conditions, 759, 1203
  for **B** and **H**, 937, 942
  for **E** and **D**, 958
Bragg, W. L., 1142
Bragg's law, 1140
Brewster, D., 1155
Brocklebank, R. W., 994
Bubble chamber, 818

Calcite, 1156, 1159
Capacitance, 741, 745
  and dielectric, 748
  and inductance, 900
  of cylindrical capacitor, 746
  of parallel-plate capacitor, 745
  of spherical capacitor, 765
Capacitative time constant, 805, 902
Capacitors, and dielectrics, 741
  as circuit elements, 802
  force on plates of, 769
  in parallel, 747
  in series, 748
Carragan, G. H., 1118
Cavendish, H., 650
Cavity, acoustic, 954
  electromagnetic, 963
Cavity radiator, 1174
CERN accelerator, 834
Charge, and matter, 655
  conservation of, 660
  in circular motion, 829
  in electric field, 675
  line of, 674
  quantization of, 654
  ring of, 673
  two kinds of, 648
Circuits, 789
  charge conservation in, 798
  current in, 792, 793
  energy conservation in, 792, 798
  potential differences in, 795
  RC, 802
Circular aperture, diffraction at, 1112
Circular polarization, 1163
Circulating charges, 829
Coaxial cable, 911, 971
  currents in, 974
  energy flow in, 987
  fields in, 974
Coblentz, 1177
Coherence, 1074
Complementarity, 1213
Compton, A. H., 1184
Compton effect, 1184
Compton shift, 1184
Concave mirror, 1041
Conduction electrons, 781, 828
  drift speed of, 781
Conductivity, 774
  electrical, 780
  thermal, 780
Conductor, 649
  insulated, 730
Conservation of charge, 660
Conservation of energy, in circuits, 792
    798
Conservative force, 710
Constant-deviation prism, 1033
Continuous spectra, 1188
Convex mirror, 1041
Corona discharge, 732
Correspondence principle, 1196
Coulomb, 652
Coulomb, C. A., 650
Coulomb's law, 657, 1192
Critical angle, 1024
Curie, P., 922
Curie temperature, 926
Curie's law, 922
Current, 770
  displacement, 962
  induced, 871
  magnetic force on a, 819
  measurement of, 800
  sense of, 772
Current balance, 848, 854
Current density, 770
Current element, 860
Current loop, **B** for, 861
  torque on, 823
Cutoff frequency, in photoelectric effect,
    1180
  in waveguide, 977
Cyclotron, 831
  resonance in, 833
Cyclotron frequency, 830
Cylindrical capacitor, 749

D, 758, 934
Damped oscillations, 950
Davisson, C. J., 1201
de Broglie, L., 1200, 1204
de Broglie wavelength, 1201
Debye, P. W., 1186
Dees, 832
Descartes, R., 1015, 1023
Deuteron, 831
Diamagnetism, 923, 937
Dichroism, 1159
Dielectric constant, 749, 750
Dielectrics, 649, 750
  and capacitors, 748
  and Gauss's law, 753
Dielectric strength, 750
Diffraction, 1035, 1199
  and Huygens' principle, 1035
  and interference, 1115, 1119
  at a circular aperture, 1112
  Fraunhofer, 1101

Diffraction, Fresnel, 1200
  from a single slit, 1102, 1105, 1108, 1109
  of matter waves, 1211
  of neutrons, 1202
  of water waves, 1037
  X-ray, 1123, 1134
Diffraction grating, 1123, 1128
Dipole, electric, *see* Electric dipole
  magnetic, *see* Magnetic dipole
Dipole antenna, 980
Dipole-dipole interaction, 942
Dispersion, 1133
Displacement, electric, 758
Displacement current, 962
Distributed elements, 954
  in acoustics, 955
  in electromagnetism, 956
Dominant mode, 977
Domains, 929, 930
Doppler, J., 1006
Doppler effect, 1006
  for light, 1007
  for sound, 1006
Double refraction, 1155
Double slit, 1104, 1115
Drift speed of electrons, 771, 781
Driscoll, R. L., 821
Duality, of light, 1183
  of matter, 1200
  wave-particle, 1203
Du Mond, J., 1189

E, *see* Electric field
Eddy currents, 891
Einstein, A., 1004, 1092, 1094, 1100, 1179, 1182, 1184
Electret, 938
Electric dipole, alternating, 720, 1148, 1167
  energy of, 679
  field due to, 681, 684, 724
  in dielectrics, 750
  induced, 720, 751, 757
  in nonuniform field, 684
  potential due to, 718
  properties of, 864
  torque on, 825
  water molecule as, 819
Electric displacement, 758, 934
Electric field, 708, 712
  and potential, 724
  definition of, 817
  energy of, 760, 912
  flux of, 686
  for a conductor, 699, 771
Electric field, for dipole, 671, 728
  for line of charge, 674, 698, 849
  for sheet of charge, 698
  in coaxial cable, 974
  in waveguide, 978
  induction of, 880
  of point charge, 670
  of ring of charge, 673
  of spherical charge, 695
Electric polarization, 756
Electric potential, *see* Potential, electric
Electric quadrupole, 683, 721
Electric vectors, table of, 760
Electromagnet, 815
Electromagnetic cavity, 955
  energy flow in, 988
Electromagnetic induction, 870
Electromagnetic oscillations, 943
  energy in, 950
  forced, 952
  quantitative, 948
  resonance in, 952
Electromagnetic shielding, 892
Electromagnetic spectrum, 993
Electromagnetic waves, 971, 982, 986, 1002
Electromotive force, 789
  gravitational analog of, 790
  induced, 871, 899
Electron, 655
  angular momentum of, 916
  e/m for, 835
  free, 649
  moment of, 917, 925
Electron microscope, 1115
Electron volt, 833
Electrostatic force, 711
Electrostatic generator, 733, 734
Elsasser, W., 1201
Emissivity, 1175
Energy, and Poynting vector, 986
  in capacitor, 760
  in circuits, 783
  in electric field, 760, 790
  in electromagnetic cavity, 986
  in *LC* oscillations, 943, 948
  in magnetic field, 790, 907
  in resonant systems, 955, 956
  quantization of, 1178
Energy level diagram, 1194
Equipotential surface, 710, 727
Essen, 1002
Ether, 1003
Euclid, 1015
Exchange coupling, [illegible], 931

Expanding universe, 1009
Extraordinary ray, 1156
Eye, sensitivity, 994

Farad, 743
Faraday, M., 647, 743, 748, 851, 852, 870, 871, 923
Faraday's law, 870, 894
  counterpart of, 959
  in traveling waves, 983
Fermat, P., 1028
Fermat's principle, 1028, 1161
Ferrimagnetism, 930
Ferroelectric, 938
Ferromagnetism, 920, 926
Fiber optics, 1025
Field, electric, *see* Electric field
  magnetic, *see* Magnetic induction
Fission, 658
Fizeau, H. L., 999
Flow field, 663
Flugge, H., 1155
Flux, of electric field, 686
  of magnetic field, 815
Flux linkages, 899
Focal length, 1058
Focal plane, 1059
Focal point, 1058
Forced oscillations, 952
Foucault, J. B. L., 1000, 1023
Fourth-power law, 1175
Franklin, B., 648
Fraunhofer diffraction, 1101
Free charge, 754
Free-electron model, 828
Fresnel, A. J., 1099, 1148
Fringes of constant thickness, 1086

Galileo, 998
Galvanometer, 824
  ballistic, 928
Gamma rays, 660
Gamow, G., 1212
Gauss, 817
Gauss's law, 686, 690, 859
  and Coulomb's law, 690
  and dielectrics, 753, 759
  applications of, 695
  for fluids, 703
  for gravitation, 704
  for magnetism, 920
Geiger, H., 700
Generator, alternating current, 895
Geometrical optics, 1035, 1037
Germer, L. H., 1201
Glancing angle, 1140
Goudsmit, S. A., 842
Grating, 1123, 1128
  dispersion of a, 1133
  resolving power of, 1132
Gravitational field, 663, 665
Gravitational force, 711
Grimaldi, F. M., 1099
Group speed, in waveguide, 978
Guide wavelength, 979

H, 815, 934, 937
Haidinger's brushes, 1152
Half-silvered mirror, 1075
Half-width, of diffraction maximum, 1111
  of resonance curve, 953
Hall, E. H., 827
Hall effect, 649, 771, 827
Hansen, 1002
Harmonic oscillator, 1178
HCl spectrum, 1190
Heaviside, O., 648
Hege, E. K., 807, 906
Heisenberg, W., 1211
Helmholtz coils, 869
Henry, G. E., 995
Henry, J., 870, 900
Henry, W. E., 923
Hertz, H., 648, 986
Hipple, J. A., 821
Hull, G. F., 995, 996
Huygens, C., 1019, 1099, 1156
Huygens' principle, 1018
  and double refraction, 1159
Hydrogen atom, 1191
Hydrogen spectrum, 1191, 1195
Hysteresis, 929

Image, 1038
  erect, 1048
  inverted, 1048
  real, 1038
  virtual, 1038
Image distance, 1043
Imes, E. S., 1190
Incident ray, 1013
Index of refraction, 1014, 1023
  principal, 1157
Induced magnetic fields, 959
Induced surface charge, 751, 754, 756
Inductance, 899
  calculation of, 901
  energy storage in, 910
  of solenoid, 901
  of toroid, 901

Induction, and relative motion, 888
  electromagnetic, 870, 875
  Faraday's law of, 872, 880
  lines of, 851
Insulated conductor, 730
Insulator, 649
Interference, 1068
  and diffraction, 1115, 1119
  double-slit, 1079
  fringes, 1070
  in thin films, 1083
Interferometer, 1090
Internal resistance, 793
Interplanar spacing, 1140
Isotopes, 658
Isotropic, 1155
Ives, H. E., 1008

Jessop, H. T., 1155
Joule heating, 784, 908
Joule's law, 785
Joyce, J., 1021, 1022
Junction theorem, 798

Kapany, N. S., 1025
Katz, J. L., 1136
Kellogg, 1190
Kelvin, Lord, 733
Kirchhoff's first rule, 798
Kirchhoff's second rule, 793

Land, E. H., 994
Laser, 1027
Lattice, 771, 781
Laue spots, 1136
Lawrence, E. O., 831
Lawton, 694
LC oscillations, 943
  analogy of to SHM, 947
Lebedev, 995
Lens, 1055
  equation for, 1057
  extended object for, 1059
  sign conventions for, 1057
Lens maker's equation, 1059
Lenz, H. F., 873
Lenz's law, 873, 883, 900
Light, 993
  and quantum physics, 1173
  and the electromagnetic spectrum, 993
  angular momentum of, 1167
  corpuscular theory of, 1023
  energy of, 995
  momentum of, 995
  moving source of, 1003
Light, polarization of, 1147
  pressure of, 995
  propagation of, 1092
  scattering of, 1167
  sensitivity of eye to, 994
  sources of, 1173
  speed of, 998, 1001, 1012
  wave theory of, 982, 1023
Line integral, 713
Lines of force, 665
Lines of induction, 815, 846
Line spectra, 1188
Loop theorem, 793
Lorentz, H. A., 648, 818, 835
Lorentz relation, 818
LR circuit, 902
LRC oscillations, 946, 950
Lumped elements, 954
  electromagnetic systems, 957
  mechanical systems, 957

Magnesia, 814
Magnetic damping, 892
Magnetic dipole, 825, 915
  B for, 863
  energy of, 826
  of bar magnet, 917
  of electron, 918
  properties of, 864
  torque on, 825
Magnetic domain, 929, 930
Magnetic field, *see* Magnetic induction
Magnetic field strength, 815, 934, 937
Magnetic force, on a current, 819
  on a moving charge, 816
Magnetic induction, 663, 814, 815
  definition of, 816
  energy and, 907
  energy density and, 910
  for a circular loop, 862
  for a dipole, 863
  for a long wire, 861
  for antiparallel wires, 867
  for a traveling wave, 982
  for current sheet, 866
  for Helmholtz coils, 869
  for parallel wires, 853
  for polygon, 868
  for rectangular loop, 869
  for solenoid, 856
  for square loop, 868
  for toroid, 859
  in a cavity, 955
  in a coaxial cable, 974
  in a waveguide, 1002

Magnetic induction, induced, 959
  time-varying, 880
Magnetic pole, 915, 920
Magnetic vectors, table of, 934
Magnetism, and matter, 915
  of nuclei, 931
  permanent, 929
Magnetite, 814
Magnetization, 922, 934
Magnetization curve, 926, 929
Magnetizing current, 961
Magnification, lateral, for lenses, 1061
    for mirrors, 1048
Malus, E. L., 1151, 1153
Marsden, E., 700
Mass-energy conversion, 660
Mass number, 658
Mass spectrometer, 842
Matter, magnetism of, 915
Matter waves, 1200
Matthias, B. T., 778
Maxwell, J. C., 647, 844, 1200
Maxwell's equations, 647
  and radiation pressure, 996
  and waves, 982
  summary of, 963
Maxwell speed, distribution, 1174
Mean free path, 781
Meter, defined in wavelengths, 1092
Michelson, A. A., 1000, 1090, 1092
Michelson-Morley experiment, 1092
Microfarad, 743
Micron, 994
Microscope, 1049, 1115
Millikan, R. A., 685, 1181, 1183, 1184
Mirror formula, 1047
Mirrors, converging, 1047
  diverging, 1047
Momentum of light, 995
Moon, radar signals from, 999
Morley, E. W., 1092
Morrison, P., 1039
Morrison, P. and E., 648
Motional emf, 876, 888
Multiloop circuits, 798
Multipoles, expansion in, 722

Néel temperature, 931
Neutron, 655
Neutron diffraction, 1202
Neutron number, 658
Newton, 1014, 1023, 1099
Newton's rings, 1088
Nichols, 995, 996
Nonreflecting glass, 1087
Nuclear atom, 697
Nuclear force, 657
Nuclear magnetism, 931
Nuclear resonance, 931
Nucleus, 655
  radius of, 656

Object, virtual, 1052
Object distance, 1043
Oersted, H. C., 647, 814, 844, 853
Ohm, 774
Ohm's law, 778, 782
Oil drop experiment, 685
Onnes, Kamerlingh, 778
Optic axis, 1159
Optical maser, 1027
Optical path, 1029, 1059, 1072
Optical reversibility, 1089
Order number, 1125
Ordinary ray, 1156
Oscillations, 943, 957
  damped, 950
  electromagnetic, 943, 947, 963
  forced, 952
  LC, 943, 947
  mechanical, 947, 957
Oscilloscope, 806

Parallel conductors, forces between, 853
Parallel-plate capacitor, 745
  energy storage in, 761
  with dielectric, 748
Paramagnetism, 920, 937
Paraxial rays, 1043
Partial derivative, 725, 984
Particles and waves, 1200
Pepinsky, R., 1139
Permeability, 937
Permeability constant, 846
Permittivity constant, 746, 848, 986
Phase angle, 948
Phase change on reflection, 1084, 1089
Phase grating, 1128
Phase speed, 978
Phasor, 1080, 1125
Photoelectric effect, 1179
Photon, 1182
Pierce, J. R., 963
Planck, M., 1174, 1177, 1179, 1182, 1188, 1191
Planck's constant, 1178, 1184, 1188, 1193
Planck's radiation formula, 1177
Plane mirror, 1013
  spherical waves on, 1038

Plane of vibration, 1147
Plane-polarization, 1147
Plane waves, reflection of, 1013, 1020, 1029
  refraction of, 1013, 1021, 1029
Plimpton, 694
Polarization, 720, 1147
  by reflection, 1153
  circular, 1163
  electric, 934
  plane of, 1147
Polarizer, 1150
Polarizing angle, 1154
Polaroid, 1150
Pole, magnetic, 915, 920
Potential, electric, 708
  and electric field strength, 711, 724
  and induction, 883, 886
  for a dipole, 719
  for a quadrupole, 721
  for charged disk, 717
  for point charges, 715, 716
  inside conductor, 730
Potential, gravitational, 711
Potential difference, 708
  and induction, 883
  in circuits, 795
  measurement of, 800
  path independence of, 710
Potential energy, electric, 722
    in electric circuits, 784
    of electric dipole, 679
    of magnetic dipole, 826
Potential gradient, 725
Potentiometer, 801
Poynting, J. H., 987
Poynting vector, 986, 995
Precession, 931
Prism, 1018, 1027
  constant deviation, 1032
Proton, 655
Ptolemy, 1032
Purcell, E. M., 931

Q, 968
Quadrupole, electric, 683, 721
Quantization, in Bohr model, 1193
  of charge, 654
  of energy, 1178
  of momentum, 1207
  of wavelength, 1204, 1207
Quantum number, 1178, 1193
Quantum physics, 1173
Quarter-wave plate, 1163
Quartz, index of, 1015

Rabi, I. I., 1190
Radiancy, 1173
Radiation, acoustic, 979
  electromagnetic, 980
Radiation pressure, 995, 1078
Radioactive decay, 658, 659
Rayleigh's criterion, 1113, 1114
RC circuits, 802
Red-shift, 1009
Reflected ray, 1014
Reflection, 1013
  at plane surface, 1038
  at spherical surface, 1035, 1041
  defects in, 1048
  diffuse, 1015
  Fermat's principle for, 1029
  Huygens' principle for, 1020
  of plane wave, 1013
  of spherical wave, 1035, 1038
  phase change on, 1084, 1089
  polarization by, 1153
  sign conventions for, 1044
  specular, 1015
  total internal, 1024
Reflection grating, 1128
Refracted ray, 1014
Refraction, 1013
  at plane surface, 1014
  at spherical surface, 1049
  double, 1155
  Fermat's principle for, 1029
  Huygens' principle for, 1021
  of plane wave, 1013
  of spherical wave, 1035
  sign conventions for, 1051
Relativistic mass, 834
Relativity, 1004
  and cyclotron, 833
  and Michelson-Morley experiment, 1092
Resistance, 774
  internal, 793
  measurement of, 812
Resistivity, 774, 775
  atomic view of, 780
  coefficient of, 775, 777
  temperature variation of, 777
Resistors, in parallel, 800
  in series, 794
Resolving power, of a grating, 1132
  of a lens, 1112
Resonance, 952
  in cyclotron, 831
  nuclear, 931
Right-hand rule, 846

Ripple tank, coherent waves in, 1075
  diffraction in, 1035
  interference in, 1070
Roemer, O., 999
Rossi, B., 997
Rowland ring, 926, 935
Rush, J. H., 998
Rutherford, E., 700
Rutherford atom, 697, 700

Sandage, A. R., 1009
Scattering, double, 1169
  of light, 1167
Schawlow, A. L., 1027
Schrödinger, E., 1204
Self-field, 676
Self-induction, 899
Shielding, electromagnetic, 892
Sign conventions, for mirrors, 1045
  for spherical surfaces, 1051
  for thin lenses, 1057
Simple harmonic motion, and *LC* oscillations, 947
Single slit, diffraction at, 1102, 1105, 1108, 1109
Slepian, Joseph, 668
Snell, W., 1015
Solar compass, 1169
Solenoid, 856, 919
Spectra, 1123, 1130
Spectral radiancy, 1173
Spectrograph, 1130
Spectrum, 1015, 1130
  absorption, 1190
  electromagnetic, 993
  line, 1188
Speed of light, 998
  and electromagnetism, 959, 986
Spin, 916
Standard cell, 802
Standard meter, 1092
Standing waves, 1203
Stationary state, 1178
Stefan-Boltzmann, constant, 1175
Stilwell, G. R., 1008
Stokes, G. G., 1089
Stopping potential, 1180
Strong, J., 1152
Superconductivity, 778
Surface charge density, 731
Surface integral, 688
Synchrotron, 834

Telescope, 1114
Television frequencies, 973
Thales of Miletus, 647
Thermistor, 779
Thin lens, 1055
Thomas, H. A., 821
Thomson, G. P., 1201
Thomson, J. J., 835
Thomson atom model, 697
Time constant, capacitative, 902
  inductive, 904
Time-of-flight spectrometer, 842
Toroid, 859
  inductance of, 901
Torque, on current loop, 823
  on electric dipole, 825
  on magnetic dipole, 826
Torsion balance, 651, 749, 996
Total internal reflection, 1024
Transmission line, 971
  acoustic, 976
Transverse wave, 1147

Ulrey, 1190
Uncertainty principle, 1210
Uniaxial crystal, 1162
Unit cell, 1137

Vacuum tube, resistance of, 779
Vacuum-tube voltmeter, 801
Van de Graaff, R. J., 733
Vector field, 663
Velocities, addition of, 1004
Venus, radar signals from, 999
Vertex, 1043
Virtual image, 1038
Volt, 708
Voltmeter, 801
von Frisch, K., 1169
von Laue, M., 1135

Water, dipole moment of, 719
Watt, 785
Wave disturbances, addition of, 1080
Wave function, 1204
  meaning of, 1208
Waveguide, 977
  cutoff frequency for, 977
  optical, 1026
Wave mechanics, 1204
Wave number, 981
Wave optics, 1035
Waves, and Maxwell's equations, 982
  electromagnetic, 971
Waves and particles, 1200
Weber, 817
Wheatstone bridge, 810

Wien, W., 890
Wien's formula, 1177
Williams, H. J., 930
Wilson, M. H., 994

X-rays, as transverse waves, 1169
  Compton effect for, 1184
  diffraction, 1031, 1134

Young, T., 1068, 1073, 1148
Young's experiment, 1027, 1069, 1104, 1115

Zacharias, J. K., 1190
Zeeman, P., 835
Zeeman effect, 842

SELECTED NUMERICAL CONSTANTS

$\pi = 3.14$

$\pi^2 = 9.87$

$e = 2.72$

$e^{-1} = 1/e = 0.368$

$\ln 2 = 0.693$

$\log e = 0.434$

$\sqrt{2} = 1.41$

$\sqrt{3} = 1.73$

$$\sin 30^\circ = \cos 60^\circ = \tfrac{1}{2} = 0.500$$

$$\cos 30^\circ = \sin 60^\circ = \sqrt{3}/2 = 0.866$$

$$\tan 30^\circ = \cot 60^\circ = \sqrt{3}/3 = 0.577$$

$$\tan 60^\circ = \cot 30^\circ = \sqrt{3} = 1.732$$

$$\sin 45^\circ = \cos 45^\circ = \sqrt{2}/2 = 0.707$$

$$\tan 45^\circ = \cot 45^\circ = 1.00$$

SELECTED CONVERSION FACTORS

(See Appendix H for a more complete list)

$180^\circ = \pi$ rad

1 radian = 57.3° = 0.159 rev

1 slug = 32.2 lb (mass) = 14.6 kg

1 kilogram = 2.21 lb (mass)

1 pound (mass) = 0.454 kg

1 atomic mass unit = 1.66×10^{-27} kg

1 meter = 39.4 in. = 3.28 ft; 1 inch = 2.54 cm

1 mile = 5280 ft = 1.61 km

1 angstrom unit = 10^{-10} meter = 0.1 mμ

1 millimicron = 10^{-9} meter

1 liter = 61.0 in.^{3}

1 ft^3 = 28.3 li

1 day = 86,400 sec

1 year = 3.16×10^7 sec = 365 days

1 mile/hr = 1.47 ft/sec = 0.447 meter/sec

1 pound = 4.45 nt; 1 newton = 0.225 lb

1 atmosphere = 29.9 in.-Hg = 76.0 cm-Hg = 1.01×10^5 nt/meter^2

1 Btu = 778 ft-lb = 252 cal = 1060 joules

1 calorie = 4.19 joules; 1 joule = 0.239 cal = 2.78×10^{-7} kw-hr

1 electron volt = 1.60×10^{-19} joule

1 horsepower = 550 ft-lb/sec = 746 watts

1 weber/meter^2 = 1 tesla = 10^4 gauss

ALSO FROM WILEY EASTERN

ENGINEERING PHYSICS Part I & II

B. V. N. Rao

Retd. Professor of Physics
Vijay College, Bangalore

During the course of his study a student of engineering finds numerous applications of the principles of phsics. In order to gain a thorough knowledge of his subiects he should therefore familiar with the basic principles of physics. To gain this knowledge he has normally to refer to a number of text books on different branches of physics and make a judicious selection of matter. This is a laborious and tedious process. These books are written to provide the student with all the information required by him at one place.

In dealing with the topics emphasis is laid more on the fundamental principles than on minute details. In physics the ability to apply principles to numerical problems is of the utmost importance. Therefore a number of worked examples are provided at the end of chapters wherever necessary. If the student carefully studies these worked examples and then solves other problems he will get better grasp of the subject than if he simply reads through the text.

CONTENTS : PART I : Motion of the Rigid Body; Simple Harmonic Motion; Gravitation; Elasticity; Surface Tension; Viscosity of Liquids; Specific Heat of Gases; Continuity of State; Thermodynamics; Radiation; Electrostatics; Current Electricity; Thermoelectricity; Photoelectric Effect.

0 85226 771 1 **Rs. 18.00**

PART II : Stationary Waves and Beats; Forced Vibrations; Acoustics of Auditoria; Ultrasonics ; Combination of Thin Prisms; Defects of Lenses; Interference of Light; Diffraction of Light; Polarisation of Light; Bohr's Theory of the Hydrogen Atom; Pauli's Exclusion Principles; Molecules and Solids; X-rays and Relativity.

0 85226 772 X **Rs. 15.00**

WILEY EASTERN LIMITED

New Delhi Bangalore Bombay Calcutta Madras Hyderabad